Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Human Resources Research Center*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

Retroactive Inhibition of Verbal Associations as a Multiple Function of Temporal Point of Interpolation and Degree of Interpolated Learning: E. J. ARCHER AND B. J. UNDERWOOD 283
Studies of Distributed Practice: III. The Influence of Stage of Practice in Serial Learning: B. J. UNDERWOOD 291
Studies of Distributed Practice: IV. The Effect of Similarity and Rate of Presentation in Verbal-Discrimination Learning: B. J. UNDERWOOD AND R. O. VITERNA 296
The Influence of Equality Judgments on the Constant Error: L. KARLIN 300
The Relationship of Emotionality to the Consummatory Response of Eating: P. S. SIEGEL AND J. J. BRANTLEY 304
Fractionation of Temporal Intervals: L. W. GREGG 307
The Effect of Sleep Prior to Learning: P. WORCHEL AND M. H. MARKS 313
Figural After-Effects: "Satiation" and Adaptation: B. H. FOX 317
Central and Peripheral Factors in the Phi Phenomenon: CAROL H. AMMONS AND J. WEITZ 327
The Effect of Reinforcement on Closely Following S-R Connections: II. Effect of Food Reward Immediately Preceding Performance of an Instrumental Conditioned Response on Extinction of that Response: M. O. NAGATY 333
Experiments on Sensory-Tonic Field Theory of Perception: I. Effect of Extraneous Stimulation on the Visual Perception of Verticality: S. WAPNER, H. WERNER, AND K. A. CHANDLER 341
Experiments on Sensory-Tonic Field Theory of Perception: II. Effect of Supported and Unsupported Tilt of the Body on the Visual Perception of Verticality: H. WERNER, S. WAPNER, AND K. A. CHANDLER 346
Experiments on Sensory-Tonic Field Theory of Perception: III. Effect of Body Rotation on the Visual Perception of Verticality: S. WAPNER, H. WERNER, AND R. B. MORANT 351
The Effect of Presenting Various Numbers of Discrete Steps on Scale Reading Accuracy: H. W. HAKE AND W. R. GARNER 358

American Psychological Association

Vol. 42 No. 5 November 1951

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.00, or $14.00 annually. Single copies are $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., Prince and Lemon Sts., Lancaster, Pa., or 1515 Massachusetts Avenue, N.W., Washington 5, D. C.

This JOURNAL publishes articles which are intended to contribute toward the development of psychology as an experimental science. Experimental work with normal human subjects is favored over work with abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental psychology or for behavior theory. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, a single experimental study or an integrated series of studies accomplished simultaneously must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors who request them are supplied with 50 free reprints without covers. However, under the new APA reprint policy no free reprints will be supplied to authors submitting manuscripts after January 1, 1952. All of the cost of the author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas.

Immediate publication can be given to articles whose authors assume the full cost of publication. Such articles must undergo the usual review by the Consulting Editors of the JOURNAL. No free reprints are given in the case of these articles.

Address all articles submitted for regular publication or for immediate publication to the editor, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

The JOURNAL adheres to the conventions concerning bibliographies, citations, preparation of tables and cuts, etc., as described by J. E. Anderson and W. L. Valentine, The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, 41, 345-376. Special attention is called to the style of numbers and abbreviations (p. 353). The term 'experimenter,' with its various forms, singular, plural, and possessive singular and plural, is abbreviated *E*, *Es*, *E*'s, and *Es*'; analogous abbreviations are used for the terms 'subject' (*S*) and 'observer' (*O*). All manuscripts must be typed double-spaced throughout, including tables, footnotes, references, and titles for figures. Tables, figures, footnotes, references, and titles for figures must be placed on separate pages at the end of the manuscript.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Journal of
Experimental Psychology

VOL. 42, NO. 5 NOVEMBER, 1951

RETROACTIVE INHIBITION OF VERBAL ASSOCIATIONS AS A MULTIPLE FUNCTION OF TEMPORAL POINT OF INTERPOLATION AND DEGREE OF INTERPOLATED LEARNING

BY E. JAMES ARCHER AND BENTON J. UNDERWOOD
Northwestern University

Previous studies which have been concerned with the relation between retroactive inhibition (RI) and temporal point of interpolation, i.e., the time between the end of original learning (OL) and the start of interpolated learning (IL), have yielded ambiguous and contradictory results (**3, 5, 6, 7, 8, 9**). Some writers have concluded that temporal point of IL is not an important variable, whereas others have concluded that it is an important variable. Furthermore, these latter writers are not in agreement as to the nature of the relationship of this dimension and RI.

The most recent report by Postman and Alper (**6**) has not clarified the problem. The results offered by these writers would indicate that not only is the temporal position of IL a variable, but an extremely complicated one as well. For several reasons we believe that the results offered do not reflect the true state of affairs. Quite possibly these unusual results were a function of several factors: (a) The investigators had only seven *Ss* in each group. Although these *Ss* were randomly assigned, there is no evidence that these groups were equal in learning ability. (b) The items were presented by reading them to *S*. Of necessity there was some variability in the rate of item presentation, and this variable is known to affect the rate of learning. (c) The items were monosyllabic nouns and adjectives whose heterogeneity may have differentially affected retention. Without further elaboration, it seems reasonable to question the results and conclusions offered by Postman and Alper.

After considering all available evidence (**1, 2, 3, 5, 6, 7, 8, 9, 11**) there is only one conclusion to draw: *the status of temporal position of IL as a variable affecting RI up to the present time is indeterminate.*

In previous studies in this area (as well as could be determined) the degree of OL, the degree of IL, and the time between OL and relearning (RL) were held constant. Only the temporal position of IL was varied. If, however, we consider RI as a function of the relative response strengths of OL and IL at the time of recall, we may note another relationship. The

response strength of IL, at the time of recall, may be varied in two ways: (a) degree of learning and (b) time between learning and recall, i.e., the temporal position. Furthermore, it was reasoned, these latter two dimensions are not functionally independent of each other since, for example, a low degree of learning late on the temporal dimension might be equivalent, for purposes of interference, to a higher degree of learning earlier on the temporal dimension. In short, an interaction of the temporal position of IL and the degree of IL seemed highly probable. Such an interaction, it was further reasoned, could account for the contradictory results of previous studies. It was the purpose of this study to vary both the degree of IL and the temporal position of IL in a design to permit the determination of this interaction. Such a design will permit the evaluation of three predictions: (a) RI is a function of degree of IL; (b) RI is a function of temporal position of IL; and (c) there is an interaction of degree of IL and temporal position of IL.

Procedure

Design.—The characteristic RI design of learn List A—learn List B—recall List A was used. The experiment was also planned for use of analysis of variance to test for interaction between the two variables. List A (OL) was always learned to the criterion of one perfect trial, and relearned after 48 hr. to the same criterion. During the 48-hr. interval, List B (IL) was presented at one of three time intervals and learned to one of three criteria. The time intervals between List A and IL of List B were 0, 24, and 48 hr.,[1] and the degrees of learning of List B were 6 of 10 syllables correct, 10 of 10 syllables correct, and 10 of 10 syllables correct plus five trials. All nine combinations of time intervals and criteria were studied, but no *S* served in more than one condition.

Counterbalancing.—To equalize effects of differences in difficulty of the two lists, one list was presented first one-half of the time and the other list was presented first the other half of the time.

To equalize any differences in retention due to sex, half of the *S*s in each condition were men and the other half were women. As best as possible, this counterbalancing of sexes was combined with that of lists. This could not be done perfectly, with ten *S*s in each condition. In effect, about one-half of the men and one-half of the women learned one list first one-half of the time.

Lists.—The lists were the same as those used by Underwood (**14**). The use of the same material will permit direct comparison with portions of these other studies. There were ten pairs of two-syllable adjectives in each list. The first members of corresponding pairs in each of the two lists were identical, conforming to the A-B, A-C paradigm. The rules for selection of particular words are to be found in Thune and Underwood (**12**).

The lists, typed on white vellum tapes, were presented on a Hull-type memory drum. Each stimulus word was presented alone for 2 sec. and then with its response word for another 2 sec. After the ten pairs had been so presented, there was an 8-sec. rest interval before the list was again presented. To reduce the possibility of serial learning, the pairs were presented in three different orders.

Practice.—No practice days were given,[2] and in the majority of cases the *S*s were naive to this type of an experiment.

Instructions.—At the first session each *S* was read a set of instructions. Since these instructions had been memorized by *E*, this reading was done in as conversational a manner as possible to maintain *S*'s interest and motivation. During the first session *S* was told how and when the words would appear and how he should try to learn the list. At the close of the first session, he was asked to try to avoid rehearsing the list or

[1] Approximately 1 min. elapsed between the end of List A and the beginning of List B under the 0-hr. condition. This interval was required to change the lists and present new instructions. The 48-hr. condition was, of course, immediately before the relearning of List A. Some variability in these intervals occurred since *S*s did not come to the laboratory at precisely the correct time.

[2] Originally, the experiment had been designed with a practice period having a 24-hr. retention interval. After the data on about the first 20 *S*s had been collected, it was found that many of the intrusions during the experimental-recall session were from the practice lists. There also was evidence that not all of the conditions were being proactively interfered with to an equal degree. For this reason the entire experiment was started again. The data for all of these preliminary *S*s were set aside.

lists which he had learned. To reduce his concern with rehearsing he was jokingly told, "Forget you were here today but don't forget to come back on Wednesday (or whenever his next appointment was scheduled)."

When learning the interpolated material, *Ss* were again instructed uniformly to learn this new list in the same way as they had learned the previous list.

At the time of recall and relearning, every effort was made through repetition of instructions and the naming of the day on which the first list had been learned, to provide an adequate set for the *S* to recall this first list.

Subjects.—All *Ss* were students in either an introductory psychology course or an experimental psychology course at Northwestern University. Because the nine cells were matched for means and variability more than 90 *Ss* were actually tested.

Results

Equality of lists.—Since the first 60 *Ss* were assigned alternately to each of the two lists, it was possible to determine if the two lists were or were not of near equal difficulty. The measure used for this comparison was the number of trials required by each *S* to attain the criterion of one perfect repetition of the list. The difference between the number of trials for the two lists was not significant ($t = .58$).

Equality of groups.—Although the *Ss* were carefully assigned, after the first 60 (see above), to obtain nine groups which were nearly equal in learning ability, a statistical check of "equality" was made. The total variance of the number of trials to learn the first list of the nine groups was analyzed. Since it is only of importance to know that the groups do not differ appreciably, the analysis is not presented here. The *F*-ratios were of the order of .01 and clearly indicated that the groups were reasonably equal in learning ability.

TABLE I

Mean Number of Correct Responses on the Recall Trial *

Degree of Interpolated Learning	Temporal Point of Interpolation		
	0 Hr.	24 Hr.	48 Hr.
6/10 correct	5.2	5.1	4.3
10/10 correct	4.9	4.6	4.3
10/10 correct + 5 trials	3.1	4.3	2.8

* The best estimate of a σ_M for these means is derived from the pooled estimate of variance, since homogeneity of variance has been demonstrated. This is .69.

TABLE II

Analysis of Variance of Correct Responses on the Recall Trial

Source of Variation	*df*	Mean Square	*F*
Time	2	5.91	1.38
Degree	2	18.31	4.26*
Order	1	1.60	
Time × degree	4	1.86	
Time × order	2	0.40	
Degree × order	2	6.53	1.52
Time × degree × order	4	4.18	
Within-group	72	4.29	

* With 2 and 72 *df*, $p_{.05} = 3.13$ and $p_{.01} = 4.92$.

Correct responses at recall as index of RI.—Since no control group was used to determine absolute amount of RI, number of correct responses may be used as an index of relative RI. RI is an inverse function of number of correct responses.

The total number of correct responses on the recall trial (first relearning trial) for each *S* was classified according to degree of IL (Degree), temporal point of IL (Time), and which one of the two lists was recalled (Order).

The mean number of correct responses on the recall trial for each group is presented in Table I. In spite of the apparently wide discrepancy between some means within rows, these differences are not significant as will be seen from the analysis of variance. The differences between means within columns, however, are significant.

After ascertaining homogeneity of variance using Bartlett's test,[3] the total variance was analyzed. The results of this analysis may be seen in Table II. The within-group variance is the proper error term. As may be seen, of the several sources of variance, only the degree of IL (Degree) is significant.[4] Since the variance attributable to different temporal points of IL (Time) and the variance attributable to interaction of different temporal points of IL and different degrees of IL (Time × Degree) are not significantly greater than the error variance, it may be concluded that only the first of our three expectations is supported, i.e., *RI is a function of the degree of IL.* The other expectations, namely, that the temporal position would be a factor and that there would be an interaction of the two variables, are clearly without support.

Analysis of covariance was applied also to these recall scores in combination with data with which they were expected to be significantly correlated. They will not be considered here since they do not change the conclusions reached through the analyses of variance.

Relearning as an index of RI.—In some previous studies the number of trials required to relearn OL has been used as an index of RI. The appropriate relearning scores from this experiment were classified according to degree of IL and temporal point of IL. Classification by order was not included since this variable is clearly not significant, and a two-classification analysis is reasonably accurate for our purposes. The total variance of these scores was analyzed. Using the within-group variance as the error term, none of the other sources of variability was found to be significant (largest $F = 1.02$, for degree of IL, which is much smaller than that required for .05 level of significance, 3.11, with 2 and 81 *df*). *At least for the materials and conditions of this experiment, we may conclude that the relearning scores do not reflect differential RI.*

TABLE III

NUMBER OF OVERT INTRUSIONS ON THE RECALL TRIAL

Degree of Interpolated Learning	Temporal Point of Interpolation			Total
	0 Hr.	24 Hr.	48 Hr.	
6/10 correct	5	6	0	11
10/10 correct	12	12	3	27
10/10 correct + 5 trials	12	4	0	16
Total	29	22	3	54

Overt intrusions at recall.—We will define an overt intrusion as the giving of a second-list response in place of a first-list response when *S* is trying to recall the first of two lists. The number of overt intrusions given on the recall trial under each condition is presented in Table III.

A graphic representation for each temporal point of IL and for all three degrees of IL combined is presented in Fig. 1. This curve is based upon the lower marginal totals of Table III. Since the X-axis is an ordered variable, we can probably accept the form of this curve as being a reasonable approximation of the "true" form.

Recognizing the danger of attributing unwarranted reliability to the data in Fig. 1, a chi-square test was applied to determine if the curve deviated significantly from a

[3] Corrected $\chi^2 = 20.70$; $df = 17$; $p_{.30} = 20.60$, $p_{.20} = 22.76$.

[4] This is only significant at about the .03 level of confidence: $F = 4.26$; $p_{.05} = 3.13$, $p_{.01} = 4.92$.

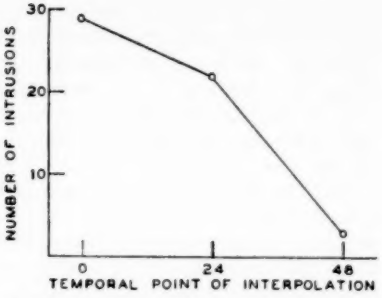

FIG. 1. Total number of intrusions on recall trial for each temporal point of interpolation

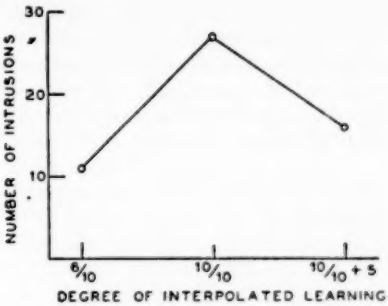

FIG. 2. Total number of intrusions on recall trial for each degree of interpolated learning

curve which is parallel to the X-axis. Since one of the three cells had only three intrusions, Yates's correction was applied. A corrected chi-square was found to be very significant.[5] These findings are to be accepted with caution in view of the unknown reliability of intrusions.

The total number of intrusions given on the recall trial for each degree of IL and for all three temporal points of IL combined is presented graphically in Fig. 2. The curve is based on the right marginal totals of Table III. Although the abscissa does not represent an ordered variable, the form of the curve can probably be accepted as a fair approximation of the "true" form. The reasoning for this is as follows: (a) the degree of IL as determined in this experiment could be related easily to the ordered variable of number of trials of presentation of IL; (b) Thune and Underwood (**12**) and Melton and Irwin (**4**), for example, have shown that for moderate degrees of OL, the number of intrusions will first increase and then decrease as the number of trials on IL is increased. On the basis of this reasoning, we may assume tentatively that the form of this curve in Fig. 2 is in agreement with previous findings.

Again, recognizing that overt intrusions may be unreliable, a chi-square test was made to determine if the curve deviated significantly from a line parallel to the X-axis. Chi-square was found to be significant,[6] which lends further support to our conclusion that the form of the curve, although a function of an unordered variable, indicates a real relationship.

DISCUSSION

Basic empirical findings

RI as a function of the degree of IL.—RI was found to increase as the degree of IL increased. This result is consistent with previous findings (**4, 12, 13**). A further consistency in the results of this experiment with previous studies (**4, 12**) is found in frequency of overt intrusions. As the degree of IL increased, overt intrusions first increased and then decreased.

RI and point of interpolation.—For the usual indices of RI (number of correct responses at recall and number of trials to relearn) RI is not a function

[5] Corrected $\chi^2 = 18.60$; $df = 2$; $p_{.01} = 9.21$, $p_{.001} = 13.82$.

[6] $\chi^2 = 7.44$; $df = 2$; $p_{.05} = 5.99$, $p_{.02} = 7.82$.

of temporal point of IL. However, a relationship between number of overt intrusions at recall and temporal point of IL was found. As the temporal point of IL increased, the number of overt intrusions decreased.

Interaction of degree of IL and point of interpolation.—From an analysis of both recall and relearning scores, we can conclude that different degrees of IL will *not* differentially affect RI at different temporal points of IL, i.e., there is no interaction of degree of IL and temporal point of IL.

Effectiveness of degree of IL

The findings of the present study are for the most part in agreement with previous investigations of the effect of degree of IL. In certain studies (**4**, **12**) RI was found to increase at a negatively accelerated rate and become asymptotic as the degree of IL increased.

It is to be noted, however, that in this study RI probably increased at a *positively* accelerated rate, but this cannot be stated with certainty since the degree of IL was not manipulated as an ordered variable, e.g., trials. However, the results indicate that with the high degree of OL which was used in this study, RI is not necessarily a negatively accelerated function of the degree of IL.

Intrusions

The finding of a sharply decreasing frequency of overt intrusions as a function of the temporal point of IL is readily explained by available concepts (**13**, **16**, **17**, **18**). The competition theory predicts that maximum frequency of intrusions will occur when two tasks are of nearly equal response strength. Owing to forgetting of OL, response strength of IL should increase (relative to OL) as (a) the temporal point of IL increases or (b) the initial degree of learning of IL increases. It is expected further that when IL is learned immediately after OL, its response strength at recall will not differ appreciably from the response strength of OL at recall. As stated before, when two response systems are about equal in strength, more intrusions will occur than when they are unequal. At the other extreme, when IL is learned immediately prior to the recall of OL, the response strength of OL is minimal, and the response strength of IL is maximal. In this case, a small number of intrusions would be expected. These were the findings of this experiment. The greatest number of intrusions occurred in the conditions in which the smallest difference in response strength at recall would be expected (IL at O hr.). The fewest number of intrusions occurred when there should have been the greatest difference in response strengths of OL and IL (IL at 48 hr.).

The finding of this experiment concerning the frequency of overt intrusions as a function of the degree of IL agrees with previous studies (**4**, **12**). When OL and IL are of about equal strength, the greatest number of intrusions should appear since competition will be greatest. As the response strengths of the two lists become increasingly unequal, fewer intrusions should occur. These were the findings of the present study.

RI and temporal position of IL

In spite of the reasoning which led us to believe that not only was the temporal position of IL a variable but that there was an interaction of temporal position and degree of IL, we found no evidence to support these beliefs. The most disconcerting aspect of these negative findings is that

we must nevertheless hold to our original beliefs or make some revisions of our concepts of forgetting. We have chosen to believe that the temporal position *is* a variable and our experiment was inadequate for its detection. It is felt that a modification of the conditions in our design would indicate the temporal position to be a factor. In brief, we believe that our negative results are only a function of the particular conditions of this experiment. This is a result which could not have been anticipated. Rather than develop an exposition for each of the possible changes, we shall only list them: (a) a lower degree of OL, (b) different total time intervals, (c) different temporal points of IL, and (d) different materials. Of these the most important probably are b and c, since different time intervals would permit a different sampling of the effects of forgetting and "spontaneous recovery" (**15**).

Summary

An experiment was performed in which degree of IL and temporal point of IL were varied simultaneously to determine the effect of each variable upon RI and their interaction. Two lists of ten paired adjectives were learned by each *S*. The OL was presented until *S* correctly anticipated all responses on a single trial. The IL was presented at 0, 24, or 48 hr. after OL and in each case was learned until *S* correctly anticipated 6 of 10 syllables correctly, 10 of 10 syllables correctly, or 10 of 10 syllables correctly *plus* five trials. The over-all retention period for OL was 48 hr.

The basic findings were:

1. As degree of IL increased, amount of RI increased. Since a control group was not used, an increase in RI was inferred from a decrease in the number of correct responses at recall.

2. Degree of IL had no effect upon relearning scores.

3. Temporal point of IL was not found to be an effective variable in determining RI.

4. There was no interaction (as determined by analysis of variance) of the two variables, degree of IL and temporal point of IL.

5. Number of overt intrusions given on recall trial was a function of the two variables in question. As degree of IL increased, number of intrusions first increased and then decreased. As time between OL and IL increased, number of intrusions sharply decreased.

The obtained results do not support the original view that there would be an interaction of the two variables, or that temporal point of interpolation is a significant variable in the determination of RI.

(Manuscript received
December 21, 1950)

References

1. Britt, S. H. Retroactive inhibition: A review of the literature. *Psychol. Bull.*, 1935, **32**, 381–440.
2. McGeoch, J. A. The influence of degree of interpolated learning upon retroactive inhibition. *Amer. J. Psychol.*, 1932, **44**, 695–708.
3. McGeoch, J. A. Studies in retroactive inhibition: II. Relationships between temporal point of interpolation, length of interval and amount of retroactive inhibition. *J. gen. Psychol.*, 1933, **9**, 44–57.
4. Melton, A. W., and Irwin, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 175–203.
5. Newman, E. B. The temporal factor in retroactive inhibition. (Abstract.) *Psychol. Bull.*, 1939, **36**, 543–544.
6. Postman, L., and Alper, Thelma G. Retroactive inhibition as a function of the

time of interpolation of the inhibitor between learning and recall. *Amer. J. Psychol.*, 1946, **59**, 439–449.

7. Robinson, E. S. Some factors determining the degree of retroaction. *Psychol. Monogr.*, 1920, **28**, Whole No. 128.
8. Sisson, E. D. Retroactive inhibition: the temporal position of interpolated activity. *J. exp. Psychol.*, 1939, **25**, 228–233.
9. Skaggs, E. B. Further studies in retroactive inhibition. *Psychol. Monogr.*, 1925, **34**, No. 161.
10. Snedecor, G. W. *Statistical methods.* Ames: Iowa State College Press, 1946.
11. Swenson, Esther J. *Retroactive inhibition: A review of the literature. Univ. Minn. Stud. Educ.*, 1941, No. 1.
12. Thune, L. E., and Underwood, B. J. Retroactive inhibition as a function of degree of interpolated learning. *J. exp. Psychol.*, 1943, **32**, 185–201.
13. Underwood, B. J. The effect of successive interpolations on retroactive inhibition and proactive inhibition. *Psychol. Monogr.*, 1945, **59**, No. 273.
14. Underwood, B. J. Retroactive and proactive inhibition after five and forty-eight hours. *J. exp. Psychol.*, 1948, **38**, 29–38.
15. Underwood, B. J. "Spontaneous recovery" of verbal associations. *J. exp. Psychol.*, 1948, **38**, 429–439.
16. Underwood, B. J. Proactive inhibition as a function of time and degree of prior learning. *J. exp. Psychol.*, 1949, **39**, 24–34.
17. Underwood, B. J. Retroactive inhibition with increased recall-time. *Amer. J. Psychol.*, 1950, **63**, 67–77.
18. Underwood, B. J. Proactive inhibition with increased recall-time. *Amer. J. Psychol.*, 1950, **63**, 594–599.

STUDIES OF DISTRIBUTED PRACTICE: III. THE INFLUENCE OF STAGE OF PRACTICE IN SERIAL LEARNING

BY BENTON J. UNDERWOOD

Northwestern University[1]

In the first study in this series (**5**) it was shown that distributed practice did not significantly facilitate the learning of serial adjectives having low intra-list similarity. With 15 sec. or with 30 sec. between each trial, learning was not faster than with 2 sec. between trials (massed). The mean differences favored learning by distribution but there was no firm basis for rejecting the null hypothesis.

A recent study by Wilson (**7**), however, showed that either 30 sec. or 60 sec. between trials *did* facilitate the learning of serial adjectives as compared with 6 sec. between trials. In examining the differences between the two studies in an effort to discover the cause of the discrepancy, two explanations seemed plausible. First, Wilson's lists had 16 items while the lists used in our study had only 14. It is known from Hovland's work (**1**) that the longer the list the greater the probability that distribution will facilitate learning. It is thus possible that a difference of two items may be sufficient to account for the findings. However, in the present study, we have chosen to investigate a second factor.

In Wilson's experiment a different group of *S*s served in each condition. This single condition was preceded by one practice session. In our study, on the other hand, *S*s served in eight sessions—two practice and six experimental sessions. Since the conditions were counterbalanced, each condition occurred after *S* had previously learned four to five lists on the average. It is apparent that our *S*s were at a much higher practice level than were those of Wilson's. On strict empirical grounds, therefore, it was reasonable to believe that stage of practice might be a variable influencing learning by distribution; more specifically, the evidence pointed to the fact that the higher the practice level the less the likelihood that distribution would enhance learning.

Although the predicted relationship between stage of practice and facilitation by distribution could be arrived at on the above empirical basis, there is also a theoretical argument for expecting the relationship. This theory, as advanced by McClelland, is in rough form. It is intended as an explanation for some of the findings on reminiscence, but may be applicable to the facts of distributed practice. McClelland notes that reminiscence is found when there is "a forced rate of performing, a massing or crowding of the material to be learned" (**2**, p. 52). He then suggests what he calls recovery from *performance inhibition* to account for reminiscence. The following quotes may serve to indicate the nature of his thinking. (His own work was with verbal-discrimination learning and the quotes refer to that, but the same argument could hold for serial learning.)

> . . . the subject is confused because he must confirm . . . his last response at the very moment when he must be making his next choice. This confusion results inevit-

[1] These experiments were made possible by a grant from The Graduate School, Northwestern University.

> ably from the fact that as soon as the reward bell rings, another pair is presented immediately. Such conflict will act to depress the subject's score below what it would be if he were not rushed.
>
> Somehow rest reduces the gap between performance and learning created by the confusion (**3**, p. 159).

As applied to serial learning the theory would say that *S* must check on the correctness of his previous anticipation and also attempt to anticipate a new item within a period of 2 sec. In a sense, *S* gets behind the drum; his attempts to correct one anticipation conflict with his attempts to give a new response. If this performance conflict does produce inhibition, the rest interval should allow recovery so that *S* gets started with the drum on the next trial.

The fact that McCelland's theory is not stated in rigorous fashion and the fact that he was concerned with reminiscence, may not lead to an unambiguous prediction concerning the influence of stage of practice on distributed training. However, the empirical fact is that *S* learns more rapidly with each successive list, the function being quite sharp for several successive lists in serial learning (**6**). It is thus possible that *S* learns to keep up with the drum more efficiently; he may learn to confirm or correct an anticipation and make a new one with greater facility. If this is the case, it would be predicted that facilitation by distributed practice would become less and less as the stage of practice increases.

We cannot be sure that McClelland will accept this deduction from his theory. Yet, in listing factors which would tend to increase conflict or inhibition in learning he says that one factor would be "inexperienced subjects who will have more difficulty with the rapid recitation procedures" (**4**, p. 243). In the present experiments the reasoning is that with each successive stage of practice *Ss* will become more and more experienced in learning and should, as a consequence, be less benefited by distributed practice.

Procedure

Materials.—Four lists of 14 two-syllable adjectives were used. Three of the lists were the same as those used in the previous experiment (**5**) and may be found in that report. These items have low intra-list similarity. They were presented at a 2-sec. rate by a Hull-type drum. On all conditions the list was learned to a criterion of one perfect trial.

Conditions.—Two groups of 24 *Ss* each (college students) served for four consecutive days. One of these groups learned a list on each day by massed practice, which consisted of a 2-sec. break between each trial. The other group learned the four lists by distributed practice, which consisted of a 30-sec. rest period after each trial. This interval was used by Wilson (**7**) and by the present writer in his earlier study (**5**). The 30-sec. interval was filled by color naming from a chart tacked above the memory drum. Rate of naming was 60 colors per minute on the first day, with an increase of ten per minute each day thereafter, as timed by a metronome.

The four lists were completely counterbalanced for both groups so that each list occurred equally often at each stage of practice.

Since it was highly desirable to introduce distribution of practice on the first day of learning, there was no completely satisfactory way of equating the two groups of *Ss* for learning ability. On each day the rest interval was not introduced until after the first anticipation trial. Thus, this trial was under identical conditions for both groups. And, while the groups show no appreciable difference in frequency of correct anticipation on this trial, computations have shown that performance on this trial and rate of learning to one perfect trial were unrelated; the correlation for both groups was near zero.

Because of the above difficulties, it must be made clear that the acceptance of the present results rests squarely on the assumption that our method of selecting *Ss* resulted in two groups of 24 *Ss* each which did not differ significantly in learning ability. In actual practice the *Ss* were rigidly assigned in an alternate fashion to the two groups. The records of four *Ss* were replaced for various reasons. Two of these were in the distributed group and two in the massed. It is therefore reasonable to believe that insofar as random assignment can result in two groups

which do not differ significantly on variables pertinent to the measured performance, this was the case in the present experiments.

Results

Trials to learn.—The basic data on number of trials to reach the criterion of one perfect trial by massing and by distribution, for the four stages of practice, are given in Table I. The values are graphed in Fig. 1.

It should first be noted that at each successive stage of practice the lists were learned more rapidly by distribution than by massing. But in no case do the differences reach a high level of reliability. Except for the first stage of practice all differences are significant between the .01 and .05 levels of confidence.

The mean differences from day to day remain almost constant. This constancy from stage to stage does *not* support the hypothesis that distribution should facilitate learning most in the early stages of practice and least in the later stages. If any change is to be noted as a function of stage of practice it is that distribution facilitates less in the early stages than in the later ones. But, the most conservative conclusion is that stage of practice is not a variable influencing learning by massing and distribution.

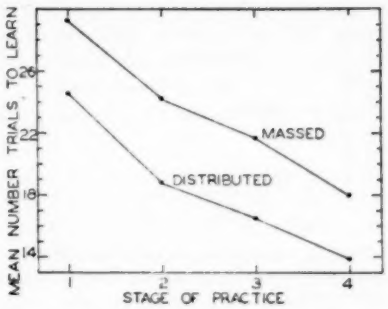

Fig. 1. The effect of stage of practice on learning by massed and by distributed practice

Errors during learning.—In the first study in this series it was found that if distribution significantly facilitated learning, more errors occurred during this distributed practice than during massed practice (5). The data of Table II and Fig. 2 show this to be true for the present conditions. At each stage of practice the mean number of errors per trial was greater for the distributed than for the massed conditions. The difference at Stage 1 is a highly reliable difference; the differences for Stages 2 and 3 fall between the .01 and .05 levels of confidence, while for Stage 4, the difference is not reliable. Evaluating the results as a whole, however, there can be little doubt that the distrib-

TABLE I

Influence of Stage of Practice on Learning by Massed and Distributed Trials

Stage of Practice	Massed		Distributed		Diff.	t*
	Mean	σ_M	Mean	σ_M		
1	29.42	1.99	24.71	2.05	4.71	1.65
2	24.17	1.99	18.92	1.21	5.25	2.25
3	21.75	1.87	16.58	1.35	5.17	2.24
4	18.04	1.58	13.88	1.01	4.16	2.21

* With 46 *df*, t's of 2.01 and 2.68 are required for the .05 and .01 levels of confidence.

TABLE II

Errors Per Trial as a Function of Stage of Practice during Learning by Massed and Distributed Trials

Stage of Practice	Massed		Distributed		Diff.	t*
	Mean	σ_M	Mean	σ_M		
1	1.05	.14	1.99	.14	−.94	4.75
2	.98	.11	1.36	.11	−.38	2.44
3	.93	.11	1.32	.14	−.39	2.19
4	.91	.11	1.22	.12	−.31	1.90

* With 46 *df*, t's of 2.01 and 2.68 are required for the .05 and .01 levels of confidence.

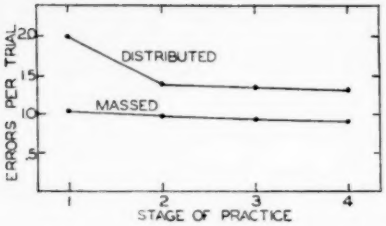

FIG. 2. The effect of stage of practice on frequency of errors in learning by massed and by distributed practice

uted conditions resulted in a greater frequency of errors than did the massed conditions. Thus, these results confirm our previous findings; number of overt errors and facilitation by distribution are related directly.

In an effort to obtain clues as to the relationship between errors and facilitation by distribution, several other analyses were made. Errors as a function of serial position were obtained for each stage of practice for massing and distribution. In no case was any new relationship apparent. The error curves as a function of serial position of the item were the same shape for the massed conditions as for the distributed; there were, of course, more errors for the distributed learning at each serial position. Furthermore, plotting the error curves as a function of stage of practice showed no change in the shapes. That is, from one stage of practice to the next, the proportion of total errors remained about the same for each serial position.

An analysis of error frequency for early and late in learning gave no clues as to the function of the errors in learning by massing and distribution. The frequency of errors on all trials needed to reach a criterion of seven correct responses was compared with the frequency of errors on trials required to go from seven to 14 correct. Proportionately no differences in these values were apparent for massing vs. distribution nor as a function of stage of practice.

It must be concluded that all we can show concerning errors is that they are more frequent when learning is by distributed practice than when learning is by massed practice. The nature, locus, and development of these errors seem to be no different for the two conditions.

DISCUSSION

The hypothesis that stage of practice and facilitation by spaced practice are inversely related has not been supported by the results of this study. Differences in learning by massed and by distributed practice remained roughly constant for four consecutive lists. It must also be concluded that if McClelland's theory (3) of performance inhibition is to be used to explain effects of distributed training, performance inhibition cannot be assigned characteristics which vary with stage of practice. The differences in learning by massed and by distributed practice were so constant for four days that it appears different asymptotes would be reached. Therefore, it cannot be argued convincingly that additional stages of practice would have produced a convergence of the two curves.

One purpose of this study was to investigate the discrepancy between Wilson's findings (7) and the first study in this series (5). Wilson found distributed practice facilitated learning of lists of 16 serial adjectives; we did not find this to be true statistically for lists of 14 adjectives, although the trends favored distributed practice. The differences in learning by massed and by distributed practice in the present study were more reliable than in our previous one, but in no case did these differences reach the .01 level of confidence. Considering the two studies together, the evidence supports the conclusion that there will be consistent but small gains in learning 14-item lists by distributed practice as compared with massed practice. Two factors might account for the fact that the differences in the present study were somewhat more reliable than in the previous one: (a) different samples of *Ss* were used; (b) because of random selection, some differences in the

ability of the two groups may have existed in the present study, i.e., the distributed group may have been slightly superior to the massed group. In view of these considerations it seems likely that Wilson's finding that distributed practice significantly enhanced learning (as compared with our marginal differences), is to be attributed to the 16-item lists which he used and not to the low level of practice of his *S*s.

In our previous (**5**) study no appreciable difference in the error frequency between massed and distributed conditions was obtained. The present data show that significantly more errors occurred under distributed than under massed conditions for the first stage of practice. By the fourth stage, the difference was unreliable, although still in the same direction as on the first stage. On strict statistical grounds, therefore, the results of our two studies are not in conflict; at comparable stages of practice the frequency of errors was not different for massed and for distributed learning in either study. Yet, the evidence of all the studies shows that when differences in error frequencies are found, the greatest frequency is with distributed practice. The present study has suggested that differences in error frequency are reduced with each successive stage of practice, but no clue has been obtained as to how these differences relate to differences in rate of learning by massed and distributed training.

Summary

Empirical evidence and theory led to the hypothesis that facilitation by distributed practice is inversely related to stage of practice. The present study was a test of this hypothesis. One group of 24 *S*s learned four serial lists of 14 adjectives each by massed practice (2 sec. between each trial), while another group with the same *N* learned the same lists by distributed practice (30 sec. between trials). Only one list was learned each session.

The results show:

1. Learning by distributed practice was consistently faster than was learning by massed training. Since the differences between the two conditions were constant for four consecutive days, the hypothesis concerning the inverse relation between facilitation by distribution and stage of practice was not supported.

2. A greater frequency of errors occurred on the distributed than on the massed conditions. The difference was highly reliable on the initial stage of practice but became unreliable by the fourth stage.

(Manuscript received
December 18, 1950)

References

1. Hovland, C. I. Experimental studies in rote-learning theory: VII. Distribution of practice with varying lengths of list. *J. exp. Psychol.*, 1940, **27**, 271–284.
2. McClelland, D. C. Studies in serial verbal discrimination learning: I. Reminiscence with two speeds of pair presentation. *J. exp. Psychol.*, 1942, **31**, 44–56.
3. McClelland, D. C. Studies in serial verbal discrimination learning: II. Retention of responses to right and wrong words in a transfer situation. *J. exp. Psychol.*, 1942, **31**, 149–162.
4. McClelland, D. C. Studies in serial verbal discrimination learning: III. The influence of difficulty on reminiscence in responses to right and wrong words. *J. exp. Psychol.*, 1943, **32**, 235–246.
5. Underwood, B. J., and Goad, D. Studies of distributed practice: I. The influence of intra-list similarity in serial learning. *J. exp. Psychol.*, 1951, **42**, 125–134.
6. Ward, L. B. Reminiscence in rote learning. *Psychol. Monogr.*, 1937, **49**, No. 220.
7. Wilson, J. T. The formation and retention of remote associations in rote learning. Unpublished Ph.D. dissertation, Stanford Univ., 1948.

STUDIES OF DISTRIBUTED PRACTICE: IV. THE EFFECT OF SIMILARITY AND RATE OF PRESENTATION IN VERBAL-DISCRIMINATION LEARNING

BY BENTON J. UNDERWOOD AND ROBERT O. VITERNA

Northwestern University[1]

An earlier study (**10**) showed that in learning serial adjectives, distribution of practice facilitated acquisition if intra-list similarity was high. It has also been demonstrated (**9, 10**) that more errors occur under distributed than under massed conditions when distribution enhances learning. These data tend to support Gibson's (**1**) hypothesis that generalization tendencies within the list recover during rests, thus resulting in more errors than would occur by massed practice. But, no clues have been offered as to how greater frequency of overt errors is related to more rapid learning.

In another report (**8**), it has been shown that intra-task similarity markedly influences the rate of learning paired-associate lists, but no differences in learning as a function of inter-trial interval were observed. Thus, even though the task was made very difficult by high intra-list similarity, distribution had no effect on learning rate.

The present experiments represent an attempt to get further information on the role played by intra-task similarity in acquisition by massed and distributed practice. The method used is that of verbal discrimination learning. This technique was employed by McClelland (**5, 6, 7**) in his studies of reminiscence, but apparently has not been used to study the influence of spaced practice. In verbal discrimination learning, *S* is presented a series of pairs of words, with one member of each pair being chosen arbitrarily by *E* as the correct response. The *S* must learn to anticipate each correct word in the series of pairs. The technique allows similarity to be varied between the two items composing a pair while at the same time keeping similarity among pairs low. Therefore, differences in learning rate can be attributed to the similarity between the two members of a pair without confounding by inter-pair similarities.

McClelland (**5**), using verbal-discrimination learning, has shown that reminiscence (as measured by relearning) will occur if a 2-sec. rate of presentation is used but not if a 4-sec. rate is employed. Hovland demonstrated that in both serial (**3**) and paired-associate (**4**) learning of nonsense syllables, increasing presentation rate will enhance the likelihood of obtaining facilitation by distributed practice (as compared with massed). In the present experiments we are investigating the effect of distributed practice in learning verbal-discrimination lists when rate of presentation and intra-pair similarity are varied.

PROCEDURE

Materials.—Each list consisted of ten pairs of two-syllable adjectives. These adjectives were taken from Haagen's lists (**2**), and chosen so that similarity among pairs was low. In Exp. I and II the similarity between the members of each pair was high (mean of 1.4 on Haagen's scale); in Exp. III and IV, the similarity was medium

[1] These studies were supported in part by a grant from The Graduate School, Northwestern University.

(mean of 2.9). Three lists of high similarity and three of medium similarity were used. The familiarity ratings for all words were high and were approximately equal for the two sets of lists. In short, if differences occur in rate of learning the two sets of lists, it is highly probable that this is a function of the differences in similarity between the two words in each pair.

The lists were presented on a Hull-type drum, using the shutter to indicate correct responses. The correct item of each pair had been chosen by *E* arbitrarily. In learning, *S* observed both words on the left side of the drum (one above the other), and was instructed to choose one or the other and call it out before the exposure period ended and the shutter lifted to show the correct word. The next pair then came into the window, and the cycle was repeated. Three different orders of pairs were used to prevent serial learning, and the positions of the two words were randomly changed from trial to trial so that *S* could not use position as a cue.

Specific conditions.—The nature of the variation in similarity has already been indicated. The rate of exposure was varied in two ways. In two experiments the two words were shown together for 2 sec., and then the correct word was shown for an additional 2 sec. before another pair moved into the window. In two other experiments, *S* had only 1 sec. in which to call out his choice when the two words were presented together, and the correct word was shown for only 1 sec.

In each of the four experiments three intertrial intervals were used; namely, 4, 15, and 60 sec. The 15- and 60-sec. intervals were filled with color naming from a board at the rate of one per second as timed by a metronome. Lists were held constant from day to day, with the conditions of distribution completely counterbalanced.

The basic conditions are summarized in Table I. It will be noted that Exp. I, III, and IV had 18 *Ss*, while II had 36. Trends which occurred in Exp. II with 18 *Ss* did not quite reach statistical significance. Therefore, another group of 18 *Ss* was run.

TABLE I

EXPERIMENTAL CONDITIONS

Exp.	*N*	Similarity	Rate of Presentation	Inter-Trial Interval
I	18	high	4 sec. (2+2)	4, 15, 60 sec.
II	36	high	2 sec. (1+1)	4, 15, 60 sec.
III	18	medium	4 sec. (2+2)	4, 15, 60 sec.
IV	18	medium	2 sec. (1+1)	4, 15, 60 sec.

The criterion of learning for each list was two successive perfect trials. The *S* was instructed to anticipate on the first trial and every trial thereafter. With only two alternatives, five items would be anticipated correctly by chance. In addition to learning the lists, *S* recalled and relearned each after 24 hr. These retention scores were uniformly high (a mean of about nine items), and since none of the conditions in any experiment was differentiated by these retention scores, they will not be presented.

All *Ss* served one practice session before starting the three experimental conditions. The list on the practice day was learned partly by massing (4 sec. between trials) and partly by distribution (30 sec. between trials). The list was first presented for five trials under massed conditions after which the drum was stopped and instructions for color naming given. Then, *S* continued learning the list by distributed practice until all items were correctly anticipated for two successive perfect trials. The practice list was made up of pairs of medium similarity and was used as the practice list for all experiments.

RESULTS

The results of all four experiments are presented in Table II. The values in each case represent the mean number of trials to reach the criterion of two successive errorless trials. The mean number of trials to learn the practice list is also shown for each group.

Influence of distribution.—Table II shows that no appreciable difference in learning is produced by distributed practice within any of the four experiments. None of the differences approaches statistical significance. It may therefore be concluded that under the conditions of similarity, rate of presentation, and distribution used in these experiments, spacing of trials neither facilitates nor retards learning.

Influence of similarity.—The influence of similarity can be determined by comparing the mean number of trials to learn under Exp. I with the comparable values for Exp. III (for slow rate of presentation), and Exp. II vs. IV for the rapid presentation rate. Since it has been shown that

TABLE II

MEAN NUMBER OF TRIALS TO LEARN

Exp.	Practice List		Inter-Trial Interval					
			4 Sec.		15 Sec.		60 Sec.	
	Mean	σ_M	Mean	σ_M	Mean	σ_M	Mean	σ_M
I	9.56	1.10	7.94	.92	7.83	.78	7.94	.74
II	13.89	1.16	12.64	1.27	12.81	1.07	13.81	1.03
III	7.12	.80	6.22	.92	6.28	.50	5.67	.48
IV	14.33	2.45	10.78	1.54	10.94	1.70	11.06	1.64

the inter-trial interval is not a pertinent variable, all conditions for a given experiment may be combined to get a mean score for each *S*.

It will be remembered that the list used on the practice day was the same for all groups. Practice-day performance scores provide a means for evaluating the equivalence of learning ability of the groups.[2] The practice-day performance scores for Exp. I and III suggest that *S*s in Exp. III were superior to those in I. Statistically the difference is not significant ($t = 1.79$), but it is great enough to suspect that differences on the experimental day may be in part a function of ability. On the experimental days *S*s who learned the medium-similarity lists (Exp. III) have consistently better performance than those who learned the high similarity lists (Exp. I). The difference, $1.84 \pm .78$, gives a t of 2.36 ($.05 > p > .01$, with 34 df).

In view of the differences on the practice day, which are in the same direction as those on the experimental days, intra-pair similarity probably has less effect than indicated by the above t. This conclusion is further supported by the comparisons between Exp. II and IV, in which the presentation rate was very rapid.

On the practice day the *S*s of Exp. II and those of Exp. IV showed almost identical mean performance ($t = .16$). On the experimental conditions learning tended to be more rapid under the medium similarity of Exp. IV. However, the difference between the experiments, 2.16 ± 1.40, is not reliable. Therefore, a conservative conclusion is that intra-pair similarity has little if any effect on rate of learning verbal discrimination lists when inter-pair similarity is low.

DISCUSSION

The results of the present experiments are uniformly negative. Distribution intervals up to 60 sec. between trials do not influence rate of learning verbal-discrimination lists of ten pairs when (a) pairs are presented at either a 2- or a 4-sec. rate, nor (b) when intra-pair similarity is high or medium. Intra-pair similarity has no appreciable effect on rate of learning.

McClelland (5) used lists of 20 pairs of adjectives having low inter- and intra-pair similarity. After *S*s had learned to a criterion of 15 out of 20 items correct, they rested 2 min. and then relearned the list to mastery. With a 4-sec. presentation rate the 2-min. interval produced no effect on relearning, whereas with a 2-sec. rate faster relearning was observed (reminiscence).

[2] The product-moment correlations between practice performance and mean performance on the experimental days are as follows: .82, .63, .91, and .83, for Exp. I to IV, respectively.

Since reminiscence and effects of distributed practice are commonly believed to have a common cause, note must be taken of the difference between McClelland's findings and ours.

Since we have shown that intra-pair similarity is a very minor factor in verbal-discrimination learning, the fact that McClelland's lists were of low similarity is probably of small consequence. The most plausible hypothesis to account for the difference between McClelland's and our results is the length of list used. When McClelland's items were presented at a 4-sec. rate the number of trials to learn was roughly comparable to the number of trials required to learn the 10-pair lists presented at a 2-sec. rate in our experiments. No benefit of rest interval(s) was observed in either experiment. But, when McClelland presented his 20-pair lists at a 2-sec. rate, the number of trials to learn was considerably greater than in any condition of the present experiments. And, it was only under such conditions that McClelland found a 2-min. rest interval to be beneficial to subsequent learning. Thus, it is probable that longer lists than used in the present experiments would produce faster learning by distributed than by massed practice.

In addition to increasing the length of the lists, the influence of *inter-pair* similarity must be investigated under varying conditions of distribution. Increasing both the length of the list and inter-pair similarity of the items should make a fairly difficult learning task, a condition which usually, but not always (**8**), tends to increase the likelihood of producing faster learning by distributed than by massed practice.

Summary

The experiments reported here used verbal-discrimination lists of ten pairs of two-syllable adjectives. Three variables were investigated in four experiments: (a) three inter-trial intervals, these being 4, 15, and 60 sec. between each trial; (b) two degrees of intra-pair similarity, and (c) two presentation rates, namely, 2 sec. per pair and 4 sec. per pair. The results show:

1. No significant influence of distributed practice under either condition of similarity or under either presentation rate.

2. No significant influence of intra-pair similarity on rate of learning.

(Manuscript received
December 18, 1950)

References

1. Gibson, Eleanor J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
2. Haagen, C. H. Synonymity, vividness, familiarity, and association value ratings of 400 pairs of common adjectives. *J. Psychol.*, 1949, **27**, 453–463.
3. Hovland, C. I. Experimental studies in rote-learning theory: III. Distribution of practice with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, **23**, 172–190.
4. Hovland, C. I. Experimental studies in rote-learning theory: VIII. Distributed practice of paired associates with varying rates of presentation. *J. exp. Psychol.*, 1949, **39**, 714–718.
5. McClelland, D. C. Studies in serial verbal discrimination learning: I. Reminiscence with two speeds of pair presentation. *J. exp. Psychol.*, 1942, **31**, 44–56.
6. McClelland, D. C. Studies in serial verbal discrimination learning: II. Retention of responses to right and wrong words in a transfer situation. *J. exp. Psychol.*, 1942, **31**, 149–162.
7. McClelland, D. C. Studies in serial verbal discrimination learning: III. The influence of difficulty on reminiscence in responses to right and wrong words. *J. exp. Psychol.*, 1943, **32**, 235–246.
8. Underwood, B. J. Studies of distributed practice: II. Learning and retention of paired-adjective lists with two levels of intra-list similarity. *J. exp. Psychol.*, 1951, **42**, 153–161.
9. Underwood, B. J. Studies of distributed practice: III. The influence of stage of practice in serial learning. *J. exp. Psychol.*, 1951, **42**, 291–295.
10. Underwood, B. J., and Goad, D. Studies of distributed practice: I. The influence of intra-list similarity in serial learning. *J. exp. Psychol.*, 1951, **42**, 125–134.

THE INFLUENCE OF EQUALITY JUDGMENTS ON THE CONSTANT ERROR

BY LAWRENCE KARLIN

New York University

The objective of the present experiment was to examine the possibility that certain artifacts are produced by the method of constant stimuli which tend to distort estimates of the constant error (CE) obtained by that method.

In a recent study using auditory stimuli, Doughty and Garner (**4**) found that the method of constant stimuli consistently yielded a smaller CE than did the method of average error. In a later study, Doughty demonstrated that the CE, calculated from a series of judgments in which the comparison stimuli are not symmetrical about the point of subjective equality (PSE), even though objectively symmetrical, is less in absolute magnitude than that of the "true" CE because of *S*'s "tendency to split his judgments 50-50 whenever presented with a difficult 2-choice situation" (**3**, p. 743).

This conclusion is basically similar to inferences made in connection with work and learning theory (**6**, **9**, **10**), to the effect that *S* tends to avoid repeating a response in two-choice situations when criteria for discrimination are minimal or absent. The present study is an attempt to test an inference based on this hypothesis. Specifically, if the preponderance of judgments in one category of judgment be assumed to give rise to a tendency not to use that particular category, then the CE resulting from the use of three categories of judgment (including "equal" judgments) should be greater than the CE resulting from the use of only two categories (no "equal" judgments permitted).

The purpose of this experiment, then, was to compare the magnitude of the CE's obtained with the use of two and three categories of judgment. Since we wished to test only the effects of forced judgments of near equality, the present study, unlike previous studies (**7**), did not use variable stimuli but rather objectively equal stimuli. This was done also in order to eliminate the possible confounding influence of context (**3**) and intraserial effects (**8**).

METHOD

Apparatus.—The method of constant stimuli was employed with intensities of tones being successively compared. A tone of constant pitch and loudness was generated by a General Radio beat-frequency oscillator, and presented to *S* through earphones. The following conditions of stimulus presentation were held constant throughout the entire experiment: stimulus frequency at 500 cps; stimulus intensity at 78 db re .0002 dynes/sq. cm.; length of intra-pair interval at 5.2 sec.; and length of inter-pair interval at 20 sec.

Procedure.—Twelve *S*s, students in elementary psychology and ignorant of the experimental purpose, participated. Each *S* attended two sessions at the same time on consecutive days. During each session, which lasted approximately 20 min., 35 judgments were made in accordance with either 2-category or 3-category types of instruction as given below. Each *S* received both types of instruction, one-half of the group beginning with the 2-category and the other half with the 3-category instructions.

In addition to the categories "louder than" and "softer than" used in the 2-category instructions, the 3-category instructions included the category "equal to." The only other difference between the two sets of instructions was that the *S*s in the 2-category situation were instructed to guess if they were not sure of a judgment.

Results And Discussion

Since the stimuli were compared in temporal succession, the CE in the present study is a time-error; the measure of the time-error used here is given by the expression: CE = 50–(100) L/L + S, where L and S refer to the number of "louder" and "softer" judgments. Consistent with previous definitions of the time-error (**11**), when L is greater than S, CE is negative. In determining the CE for the 3-category situation from the above formula, a question arises as to how to treat the equality judgments. Although Fechner and others (**5**) originally recommended dividing them equally, the procedure recommended by Culler (**2**), Guilford (**5**), and Kellogg (**7**) is to divide them proportionally between the other two categories. This latter recommendation is based partly on research (**1**) which shows that "forced" guessing yields more correct than incorrect responses. Since the rationale of the present study is intimately concerned with these questions, both of the above procedures are used in treating the results. Table I, therefore, shows three CE's for each *S*, one CE for the 2-category situation (CE_2) and two CE's for the 3-category situation, in which one of the 3-category CE's is determined from an equal division (CE_e), and the other is determined from a proportional division (CE_p) of the equality judgments.

Inspection of Table I shows that all of the CE_p values are greater in absolute magnitude than the corresponding CE_2 values. The over-all significance of these differences is indicated by the fact that the mean difference between CE_p and CE_2 is significant at the .01 level. The *t* for the mean difference obtained when signs are ignored is the more appropriate statistic and of course yields an even more significant value. On the other hand, the mean differences between CE_e and CE_2 are not statistically significant. It may be noted, however, that these differences lie in the opposite direction from differences obtained when CE_p is compared to CE_2. In general, then, CE_p shows a significant tendency to overestimate and CE_e a nonsignificant tendency to underestimate CE_2.[1]

The question that remains is which measure of the CE should be used in testing the hypothesis under investigation. Although it is believed that CE_p is the more appropriate measure,

TABLE I

Constant Errors for Two- and Three-Category Judgments

S	2-Category	3-Category	
	CE_2	CE_p*	CE_e**
1	−38.6	−50.0	−27.1
2	−41.4	−42.3	−31.3
3	32.9	38.2	18.6
4	−24.3	−26.2	−15.7
5	4.6	6.1	3.2
6	−21.4	−43.3	−18.6
7	−10.0	−15.4	−11.4
8	−21.4	−25.0	−17.1
9	−4.3	−13.6	−12.9
10	−27.1	−44.4	−22.9
11	−32.9	−46.8	−41.4
12	−21.4	−27.8	−14.3
Mean	−17.1	−24.2	−15.9
SD	19.8	24.7	19.5
Mean***	23.3	31.6	19.5
*SD****	11.8	14.0	9.6
t_{2p} = 3.1		t_{2e} = .51	
t_{2p}*** = 4.2		t_{2e}*** = 1.8	
r_{2p} = .96		r_{2e} = .92	

* Proportional distribution of equal judgments.
** Equal distribution of equal judgments.
*** Absolute values of CE's were used in computing these statistics.

[1] Although *N* is only 12, the magnitudes of the correlation coefficients shown in Table I suggest that each of the three types of CE yields comparable data as far as relationships among individuals are concerned.

the results obtained with both CE_p and CE_e are consistent with the hypothesis under consideration. This conclusion is based on the assumption that *S*'s low-confidence or "forced" judgments in the 2-category situation are governed by a bias which is similar to that which affects his other judgments in this situation (**1,5**). If this were the only factor, then CE_p would tend to *equal* CE_2 since the proportional division of the equality[2] judgments made in computing CE_p would then agree with what *S*, according to this factor, had done with his "forced" or low-confidence judgments made in the 2-category situation. On the other hand, the hypothetical tendency to avoid repeating a response will, as indicated above, attenuate CE_2 more than CE_p. Thus the joint operation of these two factors yields a CE_p which is greater than CE_2. In the case of CE_e the equal division of the equality judgments is consistent with what *S* tends to do with his low-confidence judgments solely on the basis of the tendency to avoid repeating a previous response. However, since this procedure neglects *S*'s bias which, as mentioned above, is assumed to influence both high- and low-confidence judgments similarly, CE_e will then be less than CE_2.

Apart from these considerations, the comparison of CE_p with CE_2 provides a direct test of the hypothesis free from the complicating considerations involved in using CE_e. Essentially, the present hypothesis asserts that the ratio, $L/L + S$, will change in a predictable manner from the 2- to the 3-category situation. Since the computation of CE_p is equivalent to omitting the equality judgments, just this comparison is provided. The hypothesis contains no direct specification concerning how the equality judgments should be divided but merely states that *S* will show a greater tendency to equalize the "louder" and "softer" judgments when two, rather than three, categories are used and that, therefore, as was found in the present study, the ratio, $L/L + S$, will more nearly approximate 50 per cent when two, rather than three, categories are used.

[2] It is assumed here that the equality judgments in the three-category situation are practically all low-confidence judgments. An experiment related to this assumption is in process.

Summary

When stimuli of equal objective magnitude are compared, the CE determined with the use of an equality category is greater than the CE determined with the use of only two categories, providing that the former CE is determined from a proportional division of the equality judgments. This result is discussed in terms of an hypothetical tendency to avoid repeating a previous judgment. In general, the evidence in the present study is interpreted to mean that the 3-category situation provides a more valid estimate of the CE.

(Manuscript received
January 2, 1951)

References

1. Brown, W. The judgment of differences. *Univ. Calif. Pub. Psychol.*, 1910, **1**, No. 1.
2. Culler, E. Studies in psychometric theory. *Psychol. Monogr.*, 1926, **35**, Whole No. 163, 56–137.
3. Doughty, J. M. The effect of psychophysical method and context on pitch and loudness functions. *J. exp. Psychol.*, 1949, **39**, 729–745.
4. Doughty, J. M., and Garner, W. R. Pitch characteristics of short tones. II. Pitch as a function of tonal duration. *J. exp. Psychol.*, 1948, **38**, 478–494.
5. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.

6. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
7. Kellogg, W. N. An experimental evaluation of equality judgments in psychophysics. *Arch. Psychol.*, N. Y., 1929–30, **17**, No. 112.
8. Koester, T., and Schoenfeld, W. N. The effect of context upon judgment of pitch differences. *J. exp. Psychol.*, 1946, **36**, 417–430.
9. Solomon, R. L. The influence of work on behavior. *Psychol. Bull.*, 1948, **45**, 1–40.
10. Solomon, R. L. A note on the alternation of guesses. *J. exp. Psychol.*, 1949, **39**, 322–326.
11. Woodworth, R. S. *Experimental psychology.* New York: Henry Holt, 1938.

THE RELATIONSHIP OF EMOTIONALITY TO THE CONSUMMATORY RESPONSE OF EATING[1]

BY PAUL S. SIEGEL AND JAMES J. BRANTLEY

University of Alabama

In an earlier article, Siegel and Siegel (**7**) reported an increase in the water intake of the rat immediately following intermittent faradic stimulation. This finding was discussed within the context of certain molecular theories on the nature of thirst. It was pointed out that intense emotional disturbance is accompanied by physiological changes of a kind and magnitude adequate to produce both general cellular dehydration and local dryness of the buccal cavity. The observed increase in water intake was "explained" as an adaptive response to these tissue needs. Later, Amsel and Maltzman (**2**) in an independently conceived study reported a similar empirical relationship. However, they proposed a molar explanation interpreting the finding as a confirmation of Hull's hypothesis of generalized drive (**3**). The emotional state induced by faradic stimulation was viewed as a positive contribution to the total drive level of the organism. Drinking, as a consummatory response, was thus strengthened or facilitated. Amsel and Maltzman recognize that the specific nature of the consummatory response is in no way central to this hypothesis. Any response should be augmented by the introduction of emotionality provided (a) the drive level is not already maximal, (b) the appropriate stimulus occasion for the elicitation of the response is presented, and (c) stimuli eliciting competing responses are absent. These considerations have led them directly to a differential prediction. They have reasoned that the tissue dehydration hypothesis must necessarily predict a *decrease* in the intake of dry food following emotional stimulation because of mechanical and physiological embarrassment. The generalized drive hypothesis must necessarily predict an *increase* in dry food intake (augmented consummatory response) if the above conditions are realized. Concretely, the ingestion of dry food in a familiar eating situation should be exaggerated by emotionality produced through faradic stimulation carried out in some other discrete and contrasting environment. This obviously contradicts "common sense" and the implications of the tissue dehydration hypothesis. The experiment here reported represents an attempt to test this prediction.

[1] This research was supported in part by a grant from the Research Committee of the University of Alabama.

METHOD

Subjects and apparatus.—The experiment employed 16 male albino rats of the Sherman strain. Ages ranged from 176 to 206 days. Weights ranged from 290 to 375 gm., with a mean of 328 gm.

Special dry feeding cups were utilized. The cup, described by Geo. H. Wahman Mfg. Co. as item no. LC-306, is designed to eliminate spillage.

The shock environment consisted of a rectangular compartment, 11 in. long, 6½ in. wide. and 10 in. deep, with a grid floor fashioned from brass rod, $\frac{3}{16}$ in. in diameter, spaced ½ in. apart. An inductorium served as a stimulus source. The intensity of the shock was adjusted to just yield in a test animal the typical behavioral symptoms of emotionality (mincing gait, vocalization, defecation, and urination).

The laboratory environment varied irregularly in temperature between the limits, 75° and 77° F.

Procedure.—The *S*s were handled daily until

all overt indices of resistance and disturbance had been extinguished. They were then deprived of food for 24 hr. and placed on the following schedule. With the water bottle absent, the food cup containing dry crushed Purina kibbled meal was placed in the home cage. It was removed at the end of 30 min. and the weight differential recorded. The cup was then reintroduced for another 30-min. dry feeding interval. A second measurement was made at the end of this period. To maintain normal body weight, the food cup was then returned together with the water bottle for a full hour. At the end of this period, the cup was removed. The water bottle remained for the 21½-hr. interval between daily tests. The procedure was repeated on successive evenings (6:30 to 9:00 P.M.) until the total volume of food ingested during the two dry feeding intervals had become stable. Stability was defined as three successive daily measurements yielding no statistically significant change in the mean weight of dry food consumed. Adaptation was quite rapid and the stability criterion was met by the fifth day. For Days 3, 4, and 5, the mean total intake of dry food for the two 30-min. intervals was 9.49 gm., 9.62 gm., and 9.02 gm., respectively.

The *S*s were then divided into two groups, matched on the basis of mean dry food intake during the three criterion tests. One group ($N = 8$) was arbitrarily designated as the "emotional" group; the other served as a control. On the sixth day,[2] approximately 10 min. before the first dry feeding interval, each *S* of the emotional group was removed from the home cage and placed in the shock compartment located in another room. There the animal experienced 60 sec. of intermittent faradic stimulation. An effort was made to produce no hyperactivity which might conceivably confound the results. All the details of traumatizing *S* were carried out by an *E* who had no other contact with *S*. The precautions of shocking in another experimental room and of utilizing a second *E* there were deemed advisable to minimize "carry over" from the shock situation to the eating situation. Such a generalization might conceivably produce an avoidance of food (competing response). To minimize disturbance of the eating situation, all *S*s were faradically stimulated before the introduction of the first food cup. This produced a temporal gap between emotional stimulation and food ingestion of 1 to 10 min. During the interval of punishment for the emotional group, the control group remained in the undisturbed home cage environment, i.e., the procedure did not differ from that of the preceding tests.

In the subsequent measurement of food intake during the two 30-min. intervals, except for the operation of removing and weighing the food cups between periods, the *E*'s remained out of the laboratory. This would further minimize "carry over" from shock situation to eating situation.

On the succeeding Day 7, *S*s in the control group of Day 6 were treated as emotional *S*s, i.e., shocked for 60 sec. and subsequent food intake measured.

Results and Discussion

The experimental design is "double barrelled." One may compare the food intake of each *S* immediately following emotional stimulation with his normal food intake of the preceding evening, upon the assumption that the latter constitutes a stable base line and that no uncontrolled determining factor has become operative during the interim.[3] On the other hand, one may compare the two matched groups of Day 6. Since it provides for the intrusion of uncontrolled determining factors, the latter analysis should be regarded with greater confidence.

Table I presents both analyses. The results are essentially the same in both instances. For both the first and second 30-min. measurements, the emotional *S*s yield a greater mean dry food intake. In the design of the experiment, the selection of two 30-min. intervals of measurement was dictated by an a priori consideration that the mechanical embarrassment predicted by the tissue dehydration hypothesis might be initially overrid-

[2] The *S*s were weighed in the morning. An exact specification of motivation should include the fact that at this point the *S*s had on the average been reduced to 95.2 per cent of their original body weight.

[3] During Day 6, the mean total dry food intake of the control was 9.24 gm. Comparison with the values obtained for the entire group on the three preceding adaptation days, 9.49 gm., 9.62 gm., and 9.02 gm., argues for an essential stability of this base line.

TABLE I

NORMAL DRY FOOD INTAKE AND DRY FOOD INTAKE FOLLOWING EMOTIONAL STIMULATION

Test Period	Food Intake in Grams		t	p
	Normal	Emotional		
Each S as Own Control (N = 16)				
First 30 min.	5.23	5.71	2.00	$.10 > p > .05$
Second 30 min.	3.91	4.32	1.74	$.20 > p > .10$
Total	9.14	10.03	2.94	$.02 > p > .01$
Matched Groups (N = 8)				
First 30 min.	5.23	6.03	1.95	$.10 > p > .05$
Second 30 min.	4.01	4.38	0.88	$.50 > p > .40$
Total	9.24	10.41	3.08	$.02 > p > .01$

den by intense hunger. We see, however, that the emotional *S*s eat more from the very outset. Indeed, the effect is somewhat greater during the first 30-min. interval. For the more reliable 1-hr. total, the difference justifies statistical confidence. We interpret our findings as contradictory to the tissue dehydration hypothesis and as confirming prediction from Hull's hypothesis of general drive.

At the present writing, it is impossible to make any definitive evaluation of the generalized drive hypothesis. Studies specifically designed to test the hypothesis have yielded contradictory evidence. Amsel (**1**) and Amsel and Maltzman (**2**) find both negation and support in their observations. The two positive studies of Kendler (**5**) and of Webb (**8**) are subject to alternative interpretations. Siegel (**6**) has reported negative results, but under conditions that could be regarded as unfavorable to the hypothesis. Presumably, in consideration of these findings, Hull has recently abandoned his earlier broad position in favor of the more restricted statement that: "At least some drive conditions tend partially to motivate into action habits which have been set up on the basis of different drive conditions" (**4**, p. 176).

SUMMARY

In agreement with Hull's generalized drive hypothesis, it was found that rats made emotional by faradic "teasing" ingest a greater than normal amount of dry food when subsequently placed in a familiar eating situation.

(Manuscript received
December 14, 1950)

REFERENCES

1. AMSEL, A. The combination of a primary appetitional need with primary and secondary emotionally-derived needs. *J. exp. Psychol.*, 1950, **40**, 1–14.
2. AMSEL, A., AND MALTZMAN, I. The effect upon generalized drive strength of emotionality as inferred from the level of consummatory response. *J. exp. Psychol.*, 1950, **40**, 563–569.
3. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
4. HULL, C. L. Behavior postulates and corollaries—1949. *Psychol. Rev.*, 1950, **57**, 173–180.
5. KENDLER, H. H. Drive interaction: I. Learning as a function of the simultaneous presence of the hunger and thirst drives. *J. exp. Psychol.*, 1945, **35**, 96–109.
6. SIEGEL, P. S. Alien drive, habit strength, and resistance to extinction. *J. comp. Psychol.*, 1946, **39**, 307–317.
7. SIEGEL, P. S., AND SIEGEL, H. S. The effect of emotionality on the water intake of the rat. *J. comp. physiol. Psychol.*, 1949, **42**, 12–16.
8. WEBB, W. B. The motivational aspect of an irrelevant drive in the behavior of the white rat. *J. exp. Psychol.*, 1949, **39**, 1–14.

FRACTIONATION OF TEMPORAL INTERVALS [1]

BY LEE W. GREGG

University of Wisconsin

A number of useful scales of psychological magnitudes have been constructed on the basis of data obtained by psychophysical methods. Stevens' (6) sone scale for the psychological magnitude loudness, Stevens and Volkman's (8) mel scale for pitch, and Harper and Stevens' (2) veg scale for lifted weights are representative examples. Basically, the operations involved in the construction of such scales establish a relationship between *S*'s evaluation of the magnitude of the stimuli and the conventionally determined values of the stimuli within the particular dimension. In all the above scales, the fractionation technique was used successfully to obtain the basic data. The present investigation is concerned with the application of the fractionation technique to the time dimension.[2]

The variables affecting judgments of time are numerous. Besides the more usual ones dealing with the nature of the interval and state or set of the organism, two stand out as rather unique. Both the time order error and the extent of the time range have been shown to influence temporal judgments. Stott (10) has shown that the second of a pair of 1,000-cycle tones is relatively underestimated for durations from 2.0 to 36.0 sec. Since the procedure followed in fractionation requires a successive comparison of two stimuli, the time order error should be expected to have a differential effect depending on whether the standard precedes or follows the variable stimulus. The experimental design of the present study permits an evaluation of the factor. Postman and Miller (4) found that a comparison stimulus of greater duration than stimuli of a previously judged series of durations brought about a shift in the relative judgments of the magnitudes of the durations. While this finding was obtained under very different conditions from those employed in the present study, the implication is that the generality of any time scale is limited in terms of the range of the stimuli used. A further complication which provides serious limitations from a theoretical viewpoint, at least, is the fact that different results may be obtained when a bisection or equal-sense-distance technique is employed rather than fractionation. Stevens and Davis (7) discuss this problem with reference to the sone scale, but regardless of the sensory continuum this aspect of the measurement problem is a relevant one.

In the fractionation technique, *S*'s task is that of producing a stimulus which he judges to be one half of a standard, e.g., a tone half as loud or half as high in pitch or, in this case, half as long in duration as another. The temporal intervals of this study were "filled" intervals defined by the duration of a 1,000-cycle tone. Five

[1] This report is based on a thesis submitted in partial fulfillment of the requirements for the M.S. degree. I am indebted to Dr. W. J. Brogden under whose direction the research was conducted. This investigation was supported in part by the Research Committee of the Graduate School from special funds voted by the Wisconsin State Legislature.

[2] After this article was submitted for publication, a study of subjective time by Ross and Katchmer (5) was submitted and published. However, the durations used by them for fractionation were all longer than those used in the present study.

standard durations were used. These were 400, 800, 1,600, 2,400, and 4,800 msec. The method of average error was used to obtain the judgments.

Apparatus and Procedure

Experimental design.—Each of 36 *S*s made ten judgments for each of the five standard durations. Three sequences of the standards were used: an ascending series (400, 800, 1,600, 2,400, and 4,800 msec.); a descending series (4,800, 2,400, 1,600, 800, and 400 msec.); and a random series (800, 4,800, 400, 2,400, and 1,600 msec.). Twelve *S*s were assigned at random to each sequence. Half of the *S*s in a given sequence received the standard first for five of the ten judgments at a particular duration of the standard followed by five judgments in which the variable preceded the standard. The other half received the reverse order for their comparisons.

Apparatus.—An electronic chronoscope was used to measure the durations of both the standard and the variable stimulus and to calibrate the electronic interval timer which controlled these durations. This chronoscope has been described earlier in a paper by Chernikoff and Brogden (**1**). It was calibrated to read directly in milliseconds.

The electronic interval timer has been described by Wald (**11**) and more recently by Hunter and Brown (**3**). The modified timer used in the present experiment used precision, continuously variable resistance units, trade name "Helipots." This variable resistor provides fine adjustment throughout its range by means of a single shaft rotating through 5,400 degrees—15 turns. It was essential that such provision be made since the method of average error requires *S* to directly control the variable stimulus. Unless a sufficiently great number of turns of the *S*'s control be available, motor and visual cues from the control itself invalidate the judgments. Under the conditions of operation in this experiment, the timer was accurate within one per cent.

Procedure.—The *S* sat facing a table placed against one wall of the sound-treated experimental room. On the table were a telegraph key to be used for signalling and a small black box containing the Helipot controlling the variable-duration stimulus. A pair of 2-in. loudspeakers was fastened beneath the table, one on each side to the right and left of *S*. Thus, since both of the speakers were approximately the same distance from *S*'s ears, the stimulus tones were of indefinite localization. The intensity of the tone was approximately 18 db above the threshold of hearing. A Wien bridge oscillator served as the source of the stimulus tones. The output of this oscillator was shorted through the connections of the interval timer except during the timing interval when this signal energized the loudspeakers in *S*'s room through a Hewlett-Packard attenuation pad and matching transformer. The duration of the tone was measured by the electronic chronoscope.

The experiment was carried out in two sessions. The first was a practice session in which *S* received instructions and practice on each of the durations. *S* was told that he would hear a pair of tones in rapid succession, one of which was a standard, the duration of which would remain constant, and the other of which was the variable, the duration of which he would be able to change. After listening to the first presentation of this pair of tones, he was to make a first approximation to the half-value of the standard duration by rotating the control knob in the proper direction. An arrow pointing in the clockwise direction with the word "increase," and another pointing counterclockwise with the word "decrease," were affixed to the face of the black box, which was otherwise without markings, as was the control knob. After making this first adjustment, *S* closed the telegraph key to signal *E* that he was ready to hear the tones again. The pair of tones was again presented; another adjustment was made by *S*; and he signalled to hear the tones again. *S* was told to take as long as he wished to make his decisions, and that he would be allowed to hear the pair of tones as many times as he desired. When satisfied that he had produced a tone that "sounded" half as long as the standard, *S* signalled *E* to record the setting. After each judgment, the control knob was thrown off *S*'s previous setting by *E* in a predetermined manner, by means of a remote control of *S*'s Helipot.

The instructions requested *S* to refrain from counting of any sort and cautioned him against using motor or visual cues, or such rhythmical physiological devices as breathing rate, pulse rate, or heartbeat. The *S* was also instructed to use a bracketing technique in making the successive approximations.

In the initial sessions, *S* received practice on each of the five standards in the same order as later given in the experiment proper. The experimental session was given approximately one week following the practice session. The *S* was told whether the standard or the variable stimulus was to be the first tone of the pair, but the actual values of the standards were not known to him. The interval between the stimulus tones was not constant, but varied between 1.5 and 2.0 sec. with a mean of about 1.75 sec.

Subjects.—The *S*s for this experiment were 36 students in elementary psychology at the Uni-

versity of Wisconsin. There were 19 men and 17 women. The *S*s were run in two groups of 18 which were later determined not to be significantly different.

Results

The median of the five judgments made by each *S* with the standard first and the median of the five judgments with the variable first for each standard duration were obtained. Table I presents the means of these medians for the 36 *S*s as well as the means for the combined median judgments. The standard deviations are for the combined scores. Figure 1 is a plot of the combined mean half-judgments against the durations of the standards. The vertical lines indicate the magnitudes of the standard deviations.

Bartlett's test for homogeneity of variance was applied to the combined data using the means of the median half-judgments. The significance test is between variances for the five durations of the standards. The χ^2 was found to be 145.09 based on 35 *df*; for 30 *df*, χ^2 at the .001 level is 59.70. Since the variances are not homogeneous, no comparisons of the means for the several durations of the standard were made. It is obvious from Fig. 1, however, that the particular values of the standard were such that the durations were easily discriminated.

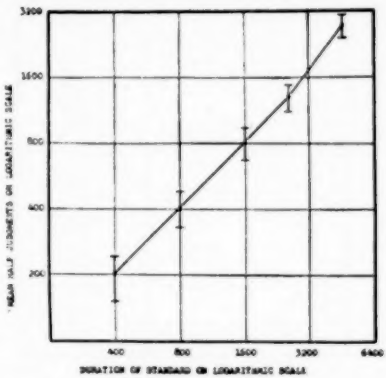

Fig. 1. Plot of mean half-judgments against the standard durations. The vertical lines represent the value of one *SD* above and below the mean for that particular duration.

For a particular duration of the standard, it is possible to test the difference between the means for the conditions of standard first and variable first. Since half of the *S*s made five judgments with the standard first followed by the variable first in the pair of tones and the other half made their judgments in the reverse order, the comparisons between the two conditions are for all 36 *S*s, the order variable being counterbalanced within the block of judgments. No significant *F*'s were found for the variance ratios for the two conditions of position of standard with respect to the variable stimulus. The separate *t*'s were carried out using a formula for related measures, and are shown in Table I. It will be seen that with the 400 msec. tone, the half-duration is judged to be longer ($p < .01$) when the standard occurs first, but with the 4,800-msec. tone the half-duration is judged to be longer ($p < .01$) when the variable occurs first. These differences are reflected in Fig. 2, which

TABLE I

Means of Median Fractionation Judgments (in Milliseconds) ($N = 36$)

Duration of Standard	Variable First	Standard First	Total		Time Error *t*
			Mean	*SD*	
400	217.8	190.1	204.0	40.5	3.29
800	424.9	391.7	408.3	77.5	1.91
1,600	839.4	779.0	809.3	140.7	2.02
2,400	1,297.1	1,259.2	1,278.2	192.8	0.90
4,800	2,521.8	2,725.0	2,623.4	329.0	3.09

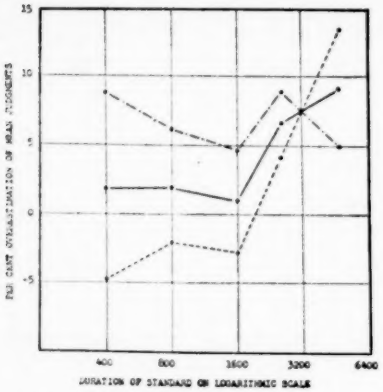

FIG. 2. Per cent overestimation of mean half-judgments. The dot-dash line connects means obtained when the standard preceded the variable stimulus; the dash line connects means obtained when the variable preceded the standard; the solid line represents all judgments. The per cent values are based on the actual value of half the standard.

shows the mean half-judgments expressed as percentages of the actual half values as a function of the duration of the standard.

DISCUSSION

The results indicate that the fractionation technique can be used to obtain a reliable time function for durations as widely disparate as those used in this study. In fact, for the range of durations covered, this function appears to be quite linear. Individual *S*s over- or underestimated the half values sometimes by as much as 50 per cent, but seemed to be rather consistent in either over- or underestimating. This is probably accounted for in terms of differing criteria from *S* to *S*. For the group as a whole, the agreement between the half judgments and the half values as determined by the conventional system of time keeping is very close. We conclude that under these conditions of measurement *S*s from this population can on the average accurately judge temporal intervals.

What are the factors making such accuracy possible? First, the experimental design was such as to permit balancing out of such variables as the sequences in which judgments for the several standards were made, the time error arising from successive comparison of standard and variable, and the order in which the judgments for standard first and variable first were made. There was no incentive for *S*s either to hurry through their judgments or prolong them. They were to report for a given period of time and expected to remain for that period. Furthermore, the bracketing technique tends to eliminate gross errors and permits *S* to make estimates as carefully as possible within the limits of discrimination. The durations of the standard were neither so long nor so short that estimation was difficult. Finally, over and above the conditions peculiar to the experimental situation is the emphasis placed on time by our society. We are called upon almost daily to make estimations of temporal intervals, and from an early age we become familiar with clocks and other time pieces and their use in scheduling our activities.

The time error is shown graphically by the broken lines of Fig. 2. The significant difference between the standard-first and the variable-first judgments at the 400-msec. duration and the significant reversal in the direction of the difference at the 4,800-msec. standard is in accord with the earlier finding of Stott (**10**). He found that for short temporal intervals, the second of a pair of tones was relatively underestimated with respect to the first. Since it is underestimated, its actual duration in time

must be longer than that of the first if the two are to be judged equal. We see that for the 400-msec. standard, a short duration, the estimated value of the variable is greater than the actual half-value when the standard is followed by the variable. Thus, the variable must have been underestimated with respect to the judged half-standard and made longer in duration to sound equal to *S*'s estimate of the half-value. Conversely, when the variable is the first tone of the pair, it follows that for short durations the variable must be made shorter in duration to "sound" equal to the shorter-appearing half-value of the standard. For longer durations, Stott found that the second of the pair of tones was relatively overestimated with respect to the first. On the basis of this, we expect that the variable should be shorter in duration when it is the second of the pair than when it is the first of the pair. This is confirmed by the direction of the significant difference for the 4,800-msec. standard.

In the fractionation method, however, we have an additional complication. At the time of the final judgment by *S*, the variable is necessarily considerably shorter than the standard. If the longer tone, the standard, precedes, the variable tone appears shorter in contrast. If the shorter tone precedes the longer, the latter appears longer. In either case, the tendency is the same, namely, to make the actual duration of the variable greater. On the basis of this, we would expect the mean half-judgments to be greater than the actual half-values.

On the basis of the time function established by these data, a scale of subjective time was constructed in the same manner as that described by Stevens (6). The procedure was as follows. One subjective time unit, one *temp*, was arbitrarily chosen as equivalent to 1,000 msec. (1 sec.). This point was first plotted on the curve of Fig. 3. From Fig. 1, the duration which was judged to be half of 1,000 msec. was found to be 505 msec. and this was assigned the value of .5 temp. The .25-temp point was obtained similarly. The value of two subjective time units was assigned to the duration judged to be twice as long as 1 sec. to give the fourth point on the curve, and so on for the fifth. The "temp" is the name suggested for the subjective time unit, and the curve connecting the above points accordingly has been designated the "temp scale."

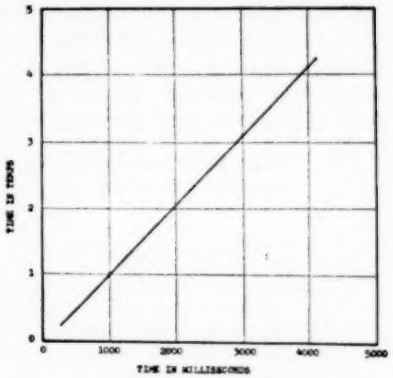

FIG. 3. The temp scale

SUMMARY

A reliable time function was established for five temporal intervals of 400, 800, 1,600, 2,400, and 4,800 msec. by the method of fractionation. Thirty-six *S*s made ten judgments each for each of the five intervals. The means of the median half-values for all *S*s were overestimations of 2.00, 2.08, 1.16, 6.52, and 9.31 per cent.

A scale of subjective time, the "temp scale," was constructed from the fractionation data, and the variables that probably influence the form of the scale were discussed.

(Manuscript received
December 11, 1950)

References

1. Chernikoff, R., and Brogden, W. J. The effect of instructions upon sensory pre-conditioning of human subjects. *J. exp. Psychol.*, 1949, **39**, 200–207.
2. Harper, R. S., and Stevens, S. S. A psychological scale of weight and a formula for its derivation. *Amer. J. Psychol.*, 1948, **61**, 343–351.
3. Hunter, T. A., and Brown, J. S. A decade-type electronic interval timer. *Amer. J. Psychol.*, 1949, **62**, 570–575.
4. Postman, L., and Miller, G. A. Anchoring of temporal judgments. *Amer. J. Psychol.*, 1945, **56**, 43–53.
5. Ross, S., and Katchmar, L. The construction of a magnitude function for short time-intervals. *Amer. J. Psychol.*, 1951, **64**, 397–401.
6. Stevens, S. S. A scale for the measurement of a psychological magnitude: loudness. *Psychol. Rev.*, 1936, **43**, 405–416.
7. Stevens, S. S., and Davis, H. *Hearing: its psychology and physiology*. New York: Wiley, 1938.
8. Stevens, S. S., and Volkman, J. The relation of pitch to frequency: a revised scale. *Amer. J. Psychol.*, 1940, **53**, 329–353.
9. Stevens, S. S., Volkman, J., and Newman, E. B. A scale for the measurement of the psychological magnitude pitch. *J. Acous. Soc. Amer.*, 1937, **8**, 185–190.
10. Stott, L. H. Time-order errors in the discrimination of short tonal durations. *J. exp. Psychol.*, 1935, **18**, 741–766.
11. Wald, S. A precision interval timer. *Electronics*, 1948, **21**, 88–89.

THE EFFECT OF SLEEP PRIOR TO LEARNING

BY PHILIP WORCHEL
University of Texas
AND
MELVIN H. MARKS
Tulane University

On the basis of their investigations of the rate of forgetting during sleep and waking, Jenkins and Dallenbach (5) proposed the hypothesis that, "forgetting is not so much a matter of the decay of old impressions and associations as it is a matter of the interference, inhibition, or obliteration of the old by the new." Although the retention curves for human *Ss* showed no greater loss after 8 hr. than after 2 hr. of sleep, there was a drop of 45 per cent in the first 2 hr. of sleep. Believing that this initial loss in recall was due to activity preceding sleep, Minami and Dallenbach (6) attempted to determine the effect of inactivity and forced activity on learning and retention in the cockroach. The results of this study suggest the possibility that the initial drop in the Jenkins-Dallenbach curve of retention during sleep may be due not to any assumed activity going on during the sleep process, but to the effect of sleep on learning capacity. As their study contains an historical section, a general review of the earlier literature is omitted.

Problem

The present study was conducted to determine (1) the effect of sleep on learning and (2) the effect on retention of sleep before and after original learning.

Procedure

Two *Ss*, M. Marks (*MM*) and J. Weiss (*JW*), served throughout the course of the experiments. Both were graduate students in psychology; neither had had any experience in memory experiments. Because dormitory sapce was limited, both *Ss* were not run concurrently. *MM* served first and then assisted with *JW*. Both *MM* and *JW* worked as *Ss* without knowledge of the specific nature or object of the problem though they may have inferred from the Jenkins-Dallenbach experiment that we were interested in the effect of sleep on retention.

A modified form of the Lippman memory drum was used. Lists of 12 nonsense syllables were prepared from Glaze's tables of three-letter syllables (2). The usual rules were followed in selecting our syllables and making up our lists.

Each list was pronounced aloud by *S* until every syllable was correctly anticipated during a single trial. If *S* made an incorrect anticipation, he was required to pronounce the correct syllable when it appeared. After a sleep interval of approximately 6¾ hr., *S* relearned the list to the same criterion. Only one list was presented at any one session. The number of trials necessary for both learning and relearning was recorded; incorrect anticipations were noted. On the experimental nights, *S* retired at 10 P.M. in the improvised dormitory set up in the laboratory. At 11:30 P.M. he was awakened by *E*. During the next 8 to 10 min., *S* washed his face with cold water in a nearby lavatory and returned to the experimental room. At this time *E* questioned *S* about his wakefulness. The *S* then seated himself in front of the apparatus which was illuminated by a 200-w. ceiling lamp. The *E* read the instructions and then stood behind *S* and recorded responses. After the final repetition, *S* retired immediately. At 6:00 A.M. *S* was again wakened by *E* and the process was repeated. On the control nights the same procedure was followed except that, instead of retiring at 10:00 P.M., *S* read or listened to the radio, and just before 11:30 P.M. prepared for bed. The learning series was then administered at 11:40 P.M.

The experimental design may be summarized as follows:

Experimental:	Sleep (1¼ hr.)	—learn—sleep (6¾ hr.)	—relearn
Control:	Awake	—learn—sleep (6¾ hr.)	—relearn

TABLE I

MEAN TRIALS AND ERRORS TO THE CRITERION OF LEARNING AND RELEARNING WITH AND WITHOUT SLEEP PRIOR TO LEARNING

Condition	N	Trials					Errors				
		Learn		Relearn		Saving	Learn		Relearn		Saving
		Mean	σ_M	Mean	σ_M	%	Mean	σ_M	Mean	σ_M	%
Subject *MM*											
Prior sleep	10	18.0	1.10	5.9	0.47	67	114.9	10.80	27.4	1.80	76
Prior waking	10	12.6	0.73	4.9	0.50	60	80.1	4.57	19.2	2.23	76
Difference	10	5.4	1.11				34.8	11.80			
Subject *JW*											
Prior sleep	9	12.4	1.27	3.7	0.53	70	73.6	7.84	17.7	2.16	76
Prior waking	9	8.1	0.71	3.6	0.78	56	48.4	4.73	16.2	2.79	67
Difference	9	4.3	0.62				25.1	5.11			

The experimental and control sessions were given in a counterbalanced order: ABBA, BAAB, etc. The order was always unknown to *S*. Both *S*s after each session reported that they remembered being awakened for the learning experiments. In order to accustom the *S*s to the experimental conditions, for five days prior to the experiment proper they slept in the laboratory dormitory. They retired at 10:00 P.M. each night, after having learned a practice series of nonsense syllables. During the experiment, sessions were conducted daily with the exception of Saturdays. The *S*s did not object to the routine, nor at any time were they indisposed. Twenty series (10 control, 10 experimental) were given to *MM*, and 18 (9 control, 9 experimental) to *JW*.[1]

RESULTS

The number of trials required to reach criterion on each of the five practice days was 25, 16, 11, 16, and 14 for *MM*, and 20, 13, 16, 15, and 8 for *JW*. Since each *S* served as his own control, and the control sessions were given in counterbalanced order with the experimental sessions, it was not considered necessary to continue the practice until a definite plateau was reached.

[1] *JW* served in 20 series, but one pair was excluded from the statistical analysis. The *S* admitted that he had fallen asleep prior to reporting for one of the control sessions.

Table I gives the learning and relearning measures for each *S* under the sleep and waking conditions prior to original learning. The number of incorrect anticipations required to reach the criterion includes overt errors and failures to respond.

The mean differences between the experimental and control relearning scores were calculated from the differences between paired measures (each experimental session versus a corresponding control session). Neither of the obtained differences between repetitions for relearning is significant, and only the difference between errors for *MM* is significant at the .05 level of confidence. It is apparent that a period of sleep before learning does not lead to better retention under the conditions of our experiment.

The second problem of this study was to determine the effect of sleep on learning. As shown in Table I, more trials and errors were required to bring learning to the criterion when sleep preceded learning than when normal "waking" activity preceded learning. All of the differences are significant at the .01 level of confidence.

The number of incorrect anticipations for each of the 12 serial positions in both experimental and control sessions for the two *Ss* shows the usual bow-shaped distribution. The introduction of sleep prior to learning had no effect on the form of the distribution.

Discussion

Ever since experimental work on forgetting as a function of intervening activity was introduced by Müller and Pilzecker (**7**), efforts have been directed towards comparing the effects of various degrees, kinds, and amounts of intervening activities on retention. It has been assumed that the effect, if any, was on "retention." The investigators have failed to consider the possibility that conditions at the time of recall or relearning may have affected performance. Before any conclusions can be drawn on the effect of activity on "retention" it is necessary to fractionate out that portion of the decrement due to the effect of the activity on learning ability.

Hunter (**4**), for example, found that cockroaches, when immobilized by cold immediately after learning an avoidance response to darkness, took longer to relearn this task than controls that were held at normal temperatures and permitted normal activity during the learning-relearning interval. It is necessary to determine the effect of immobilization by cold on learning so that we may determine what effect, if any, cold has on retention. Again, all investigations concerning the effect of drugs given some time between learning and a measure of retention should include a determination of the effect of these drugs on learning ability (**8, 9, 10, 13**).

The present study shows that 1½ hr. of sleep prior to learning significantly impairs that learning with respect to learning of equivalent material preceded by normal waking activity. It may be that some physiological process taking place either during sleep or upon awakening just prior to learning impairs the learning performance. The initial decrements observed in "retention" curves under conditions of interpolated sleep, may be due not to a decrease in the number or strength of the associations established during learning, but rather to the effect of the prior sleep in changing the learning capacity at the time of relearning.

It might be contended that the effect of sleep on learning was due to the "sleepiness" of the *Ss* when they were required to learn immediately after sleep. The results of one trial with *JW*, which was not included in the tables, are suggestive. The *S* came to the laboratory at 11:30 P.M. for a control session. After learning the list of syllables, he remarked that he had fallen asleep at about 9:30 and did not awaken till 11:00 P.M. He drove his car to the laboratory. One-half hour intervened between his awakening and the learning. The number of trials required to attain the criterion in this one session was 13. This score is far above the mean (8.1) of the control sessions, and definitely shows the impairment in learning ability due to sleep.

The results show that activity for 1½ hr. before learning is not responsible for the drop in retention. All activity in learning prior to that may, however, affect retention. The hypothesis, therefore, that the drop in retention following sleep may be due to activity prior to learning is not tenable. We must seek elsewhere for an explanation of forgetting when no overt activity is interpolated. Factors other than retroactive inhibition may be responsible for loss of memory. It is interesting, however, that so few trials are required for relearning after 6¼ hr. of sleep. *JW* required an average of only 3.7 and 3.6 trials for relearning in the control and experimental sessions respectively. Also, though both *Ss* required more practice on the original learning in the experimental sessions than in the control sessions, the number of trials required for relearning was approximately the same.

One possible explanation for the better learning ability following a period of normal waking activity than after a period of sleep may be that *S*, having been awake, would be more "warmed up" than an *S* just awakened. Thune (**12**) has shown that the rate with which a list of paired associates is learned bears a direct relation to the number of warm-up trials completed 10 min. previously. The "warm-up" effects of activity prior to learning must be investigated more systematically, however, in order to determine the nature and degree of activity in relation to "warm-up" effects since others (**1, 3, 6, 11**) have found a detrimental effect of activity on learning capacity in animals.

Summary and Conclusions

Two *Ss* learned lists of 12 nonsense syllables to a criterion of one perfect repetition with and without 1½ hr. of

TABLE I

MEAN TRIALS AND ERRORS TO THE CRITERION OF LEARNING AND RELEARNING WITH AND WITHOUT SLEEP PRIOR TO LEARNING

Condition	N	Trials					Errors				
		Learn		Relearn		Saving	Learn		Relearn		Saving
		Mean	σ_M	Mean	σ_M	%	Mean	σ_M	Mean	σ_M	%
Subject *MM*											
Prior sleep	10	18.0	1.10	5.9	0.47	67	114.9	10.80	27.4	1.80	76
Prior waking	10	12.6	0.73	4.9	0.50	60	80.1	4.57	19.2	2.23	76
Difference	10	5.4	1.11				34.8	11.80			
Subject *JW*											
Prior sleep	9	12.4	1.27	3.7	0.53	70	73.6	7.84	17.7	2.16	76
Prior waking	9	8.1	0.71	3.6	0.78	56	48.4	4.73	16.2	2.79	67
Difference	9	4.3	0.62				25.1	5.11			

The experimental and control sessions were given in a counterbalanced order: ABBA, BAAB, etc. The order was always unknown to *S*. Both *S*s after each session reported that they remembered being awakened for the learning experiments. In order to accustom the *S*s to the experimental conditions, for five days prior to the experiment proper they slept in the laboratory dormitory. They retired at 10:00 P.M. each night, after having learned a practice series of nonsense syllables. During the experiment, sessions were conducted daily with the exception of Saturdays. The *S*s did not object to the routine, nor at any time were they indisposed. Twenty series (10 control, 10 experimental) were given to *MM*, and 18 (9 control, 9 experimental) to *JW*.[1]

RESULTS

The number of trials required to reach criterion on each of the five practice days was 25, 16, 11, 16, and 14 for *MM*, and 20, 13, 16, 15, and 8 for *JW*. Since each *S* served as his own control, and the control sessions were given in counterbalanced order with the experimental sessions, it was not considered necessary to continue the practice until a definite plateau was reached.

[1] *JW* served in 20 series, but one pair was excluded from the statistical analysis. The *S* admitted that he had fallen asleep prior to reporting for one of the control sessions.

Table I gives the learning and relearning measures for each *S* under the sleep and waking conditions prior to original learning. The number of incorrect anticipations required to reach the criterion includes overt errors and failures to respond.

The mean differences between the experimental and control relearning scores were calculated from the differences between paired measures (each experimental session versus a corresponding control session). Neither of the obtained differences between repetitions for relearning is significant, and only the difference between errors for *MM* is significant at the .05 level of confidence. It is apparent that a period of sleep before learning does not lead to better retention under the conditions of our experiment.

The second problem of this study was to determine the effect of sleep on learning. As shown in Table I, more trials and errors were required to bring learning to the criterion when sleep preceded learning than when normal "waking" activity preceded learning. All of the differences are significant at the .01 level of confidence.

The number of incorrect anticipations for each of the 12 serial positions in both experimental and control sessions for the two *Ss* shows the usual bow-shaped distribution. The introduction of sleep prior to learning had no effect on the form of the distribution.

Discussion

Ever since experimental work on forgetting as a function of intervening activity was introduced by Müller and Pilzecker (**7**), efforts have been directed towards comparing the effects of various degrees, kinds, and amounts of intervening activities on retention. It has been assumed that the effect, if any, was on "retention." The investigators have failed to consider the possibility that conditions at the time of recall or relearning may have affected performance. Before any conclusions can be drawn on the effect of activity on "retention" it is necessary to fractionate out that portion of the decrement due to the effect of the activity on learning ability.

Hunter (**4**), for example, found that cockroaches, when immobilized by cold immediately after learning an avoidance response to darkness, took longer to relearn this task than controls that were held at normal temperatures and permitted normal activity during the learning-relearning interval. It is necessary to determine the effect of immobilization by cold on learning so that we may determine what effect, if any, cold has on retention. Again, all investigations concerning the effect of drugs given some time between learning and a measure of retention should include a determination of the effect of these drugs on learning ability (**8, 9, 10, 13**).

The present study shows that 1½ hr. of sleep prior to learning significantly impairs that learning with respect to learning of equivalent material preceded by normal waking activity. It may be that some physiological process taking place either during sleep or upon awakening just prior to learning impairs the learning performance. The initial decrements observed in "retention" curves under conditions of interpolated sleep, may be due not to a decrease in the number or strength of the associations established during learning, but rather to the effect of the prior sleep in changing the learning capacity at the time of relearning.

It might be contended that the effect of sleep on learning was due to the "sleepiness" of the *Ss* when they were required to learn immediately after sleep. The results of one trial with *JW*, which was not included in the tables, are suggestive. The *S* came to the laboratory at 11:30 P.M. for a control session. After learning the list of syllables, he remarked that he had fallen asleep at about 9:30 and did not awaken till 11:00 P.M. He drove his car to the laboratory. One-half hour intervened between his awakening and the learning. The number of trials required to attain the criterion in this one session was 13. This score is far above the mean (8.1) of the control sessions, and definitely shows the impairment in learning ability due to sleep.

The results show that activity for 1½ hr. before learning is not responsible for the drop in retention. All activity in learning prior to that may, however, affect retention. The hypothesis, therefore, that the drop in retention following sleep may be due to activity prior to learning is not tenable. We must seek elsewhere for an explanation of forgetting when no overt activity is interpolated. Factors other than retroactive inhibition may be responsible for loss of memory. It is interesting, however, that so few trials are required for relearning after 6¼ hr. of sleep. *JW* required an average of only 3.7 and 3.6 trials for relearning in the control and experimental sessions respectively. Also, though both *Ss* required more practice on the original learning in the experimental sessions than in the control sessions, the number of trials required for relearning was approximately the same.

One possible explanation for the better learning ability following a period of normal waking activity than after a period of sleep may be that *S*, having been awake, would be more "warmed up" than an *S* just awakened. Thune (**12**) has shown that the rate with which a list of paired associates is learned bears a direct relation to the number of warm-up trials completed 10 min. previously. The "warm-up" effects of activity prior to learning must be investigated more systematically, however, in order to determine the nature and degree of activity in relation to "warm-up" effects since others (**1, 3, 6, 11**) have found a detrimental effect of activity on learning capacity in animals.

Summary and Conclusions

Two *Ss* learned lists of 12 nonsense syllables to a criterion of one perfect repetition with and without 1½ hr. of

prior sleep. Both *S*s relearned all lists to same criterion after 6¼ hr. of sleep. One *S* was given 20 sessions; the other 18. The order of presentation of experimental and control session was counterbalanced.

Under the conditions of this study the following conclusions were reached:

1. A period of 1½ hr. sleep preceding a learning session significantly impairs learning ability.

2. Relearning after a period of approximately 6¼ hr. of sleep is unaffected by normal waking activity or sleep for 1½ hr. preceding learning.

(Manuscript received
December 27, 1950)

References

1. Gagné, R. M., and Graham, C. H. The effect of an "emotional state" on the initial stages of acquisition in a conditioned operant response. *Proc. Nat. Acad. Sci.*, 1940, **26**, 297–300.
2. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
3. Higginson, G. D. The after-effects of certain emotional situations upon maze learning among white rats. *J. comp. Psychol.*, 1930, **10**, 1–10.
4. Hunter, W. S. The effect of inactivity produced by cold upon learning and retention in the cockroach, *Blatella Germanica*. *J. genet. Psychol.*, 1932, **41**, 254.
5. Jenkins, J. G., and Dallenbach, K. M. Obliviscence during sleep and waking. *Amer. J. Psychol.*, 1924, **35**, 605–612.
6. Minami, H., and Dallenbach, K. M. The effect of activity upon learning and retention in the cockroach, *Periplaneta americana*. *Amer. J. Psychol.*, 1946, **59**, 1–58.
7. Müller, G. E., and Pilzecker, A. Experimentelle Beiträge zur Lehre vom Gedächtnis. *Z. Psychol.*, 1900, Ergbd. **1**, 1–300.
8. Plath, O. E. Do anesthetized bees lose their memory? *Amer. Natur.*, 1924, **58**, 162–166.
9. Rösch-Berger, K. Das Gedächtnis der Bienen nach der Narkose. *Z. vergl. Physiol.*, 1933, **18**, 474–480.
10. Russell, R. W., and Hunter, W. S. The effect of inactivity produced by sodium amytal on the retention of the maze habit in albino rats. *J. exp. Psychol.*, 1937, **20**, 426–436.
11. Szymanski, J. S. Abhandlungen zum Aufbau der Lehre von den Handlungen der Tiere. *Arch. ges. Physiol.*, 1918, **170**, 159–177.
12. Thune, L. E. The effect of different types of preliminary activities on subsequent learning of paired-associate material. *J. exp. Psychol.*, 1950, **40**, 423–438.
13. Tirala, L. Ueber den Einfluss der Äthernarkose auf die Heimkehrfähigkeit der Bienen. *Arch. exp. Pathol. Pharmak.*, 1923, **97**, 433–440.
14. Van Ormer, E. B. Retention after intervals of sleep and waking. *Arch. Psychol., N. Y.*, 1932, **21**, No. 137.

FIGURAL AFTER-EFFECTS: "SATIATION" AND ADAPTATION [1]

BY BERNARD H. FOX

The George Washington University

The phenomena of figural after-effects are well known by this time. However, a brief recapitulation of certain facts is important for the purposes of this paper. Gibson (**4**) described the apparent tilt of a vertical line after the inspection of another line inclined at a small angle to the vertical. He postulated that the effect was due to the tendency of the inclined line to adapt toward the norm of verticality, a phenomenon which others (**16, 17**), in a slightly different context have described. The tilted appearance of the vertical line, then, was the result of this adaptation toward a norm. Köhler and Wallach (**12**), on the other hand, declared that the effects of Gibson (**4, 5**) and others (**1, 6**) were all manifestations of a different phenomenon which they called "satiation." They maintained that the effects of satiation are such that the inspection of a figure with a well-defined border will cause a later fixated figure placed near the position of the first to be phenomenally displaced away from the boundary of the first figure. Thus, they maintained, the phenomena observed by Gibson should be explained by satiation, and Gibson's interpretation involving adaptation toward a norm should not be applied in the cases he described. Neither investigator apparently gave serious thought to the hypothesis that both phenomena might contribute to the observed displacements.

Köhler and Wallach (**12**) derived quantitative data on the amount of figural after-effect (in this case amount of apparent displacement) as a function of the distance between the first inspected figure, called the *I* figure, and the later fixated figure, called the *T* figure. They found that with medium distances (about $\frac{1}{4}$ in.) the figural after-effect was larger than that found with either greater or smaller distances between *I* and *T* figures. With still greater distances the effect declined. They called this phenomenon the "distance paradox."

In the first experiment of this investigation one of the Köhler-Wallach *I* figures (**12**), critical for their theory, was selected for study in modified form. An attempt was made to investigate some of the above distance relationships, but some of the results were inconsistent with certain theoretical implications made by Köhler and Wallach (**12**) about the strength of figural after-effects under light-dark or dark-light figure and ground relations. In one case the *amount* of apparent displacement differed from that predicted. In another, *direction* differed from that predicted. The suggestion was made that figural after-effects were operating due to adaptation to the norm in Gibson's sense (**4**) as well as to satiation in the sense of Köhler and Wallach (**12**).[2]

Experiment II was performed to

[1] Experiment I of this study was done as part of a doctoral dissertation at the University of Rochester in 1949 (**2**). Experiment II was done after the completion of the dissertation. The writer wishes to acknowledge his gratitude to Dr. G. R. Wendt for his suggestions and helpfulness during the course of this study.

[2] The writer is grateful to Dr. G. R. Wendt for suggesting this hypothesis, based on his own unpublished work.

test this hypothesis. By presenting an *I* figure symmetrically placed with respect to the fixation point any tendency toward adaptation to a norm was eliminated. The results of the two experiments differed in certain crucial aspects, those of the second experiment apparently showing only the effects of satiation. In Experiment II the "distance paradox" of Köhler and Wallach was verified insofar as it applied to satiation.

Experiment I

Apparatus

In general the apparatus consisted of two adjacent black boxes containing panes of glass. These panes, to which were affixed the black tape stimulus figures, were illuminated from behind through a pane of flashed opal glass. One box contained the pane holding the *I* figure which was immovable. The other box contained three panes of glass, to each of which was affixed one element of the *T* figure. The fixation cross was fixed to a stationary pane. Each of the other two elements, the dots, was fixed to a separate pane, both pieces of glass being capable of simultaneous vertical movement in opposite directions, so that when one dot moved up, the other moved a like distance down. The *S*'s settings were determined from a vernier scale, reading to the nearest .1 mm., attached to the edge of one movable pane.

Lighting.—Each of the boxes contained the same lighting arrangement. Across the rear panel of each box nine 6-v. lamps were arranged in a square array, wired in parallel and supplied with a.c. from a transformer. In both boxes all parts of the flashed opal panes were illuminated about equally, as measured by a Luckiesh-Taylor brightness meter from *S*'s head position. Average brightness, derived from ten readings each, was obtained for each of the following: stimulus figures, 1.804 ft.-lamberts; ground of *I* figure, 24.76 ft.-lamberts; ground of *T* figure, 20.40 ft.-lamberts. The difference between the last two was due to the presence of two extra panes of glass in the *T*-figure box, each having a transmission of about .90.

The stimulus figures.—Figure 1 shows the different *I* figures—rectangles and fixation cross—together with the *T* figures—dots and cross—used throughout. Not shown is X_X, a control *I* figure consisting only of a fixation cross in the square light field. It must be remembered that the *I* figure was inspected first, then the *T* figure was adjusted. The representation in Fig. 1 is given for convenience. The dots were ¼ in. in diameter. The total field was 11¼ in. square, and if drawn to scale, would appear larger than the field for each figure shown in the diagram. Köhler and Wallach had only one rectangle in the *I* figure and used small squares as the *T* figure.

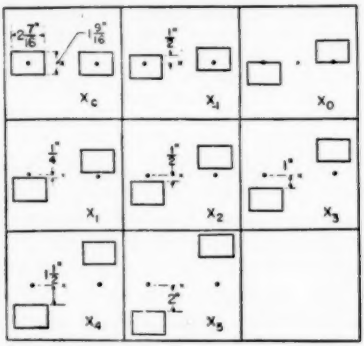

Fig. 1. Stimuli used in Exp. I. Both the *I* figures, rectangles and cross, and the *T* figure, dots and cross, were solid black.

Angular relationships.—The *S* viewed the stimulus figures as reflections in a plate glass mirror. At each session *S* was placed so that he could see his eye reflected at the position of the cross in both *I* and *T* figures, insuring that the rays of light from eye and cross were perpendicular to the glass pane at the cross. The visual distance from eye to each cross was 80 in. Each rectangle of the *I* figures subtended a visual angle of 1°8′ in the vertical dimension (Köhler and Wallach's subtended 33′).

A wooden head rest was used, adjustable for height, head width, and lateral position. An adjustable stool was provided for *S*s of different height.

Procedure

The method of average error was used throughout. At the end of a 1½-min. inspection of the *I* figure, *E* snapped a switch that simultaneously turned off the *I*-figure field and turned on the *T*-figure field. By turning a knob *S* then adjusted the two dots to the apparent horizontal, while fixating the cross between them. The *E* continuously stressed the importance of fixating the cross while inspecting and while adjusting the dots. When *S*'s hand left the setting knob, *E* read the time of adjustment from a stopwatch, read the setting from the vernier, and turned the switch back to the *I* figure. The watch was

started again from zero at that point. Depending on the length of time for setting, the total time from the end of one inspection to the beginning of the next was on the order of 15 sec. Ten trials were run for a session, during which only one *I* figure was used.

During the inspection period *E* turned *S*'s setting knob from his last setting so as to produce about an 8-mm. relative movement of the dots. The direction of this displacement alternated with each trial, but for all sessions the same order of alternation was used for the same *S*. The direction of setting for the first session was random for any *S* on his first trial.

Controls

Three questions were to be answered in setting up the controls. First, does the effect of a control stimulus which is a member of the series being investigated (that is, which has rectangles in its field—X_C in Fig. 1) differ from that of a control stimulus which has only a fixation point in its field (X_X)? To answer this question six *Ss* were given one session with X_C and the other six with X_X.

Second, are the controls reliable? Will a repetition of a control yield a similar mean setting? To answer this question all 12 *Ss* were tested with X_C in the second session.

Third, does the presentation of a series of experimental stimuli have any effect on the control if it is taken after the series? After a series of six experimental stimuli were presented, all *Ss* were retested with X_C.

A system of rotation of experimental sessions was devised. By this means the sequence of stimuli (except X_1 and the controls)[3] was such that every stimulus occupied every position in the sequence twice only and every stimulus followed every other stimulus twice only, in order that the influences of position and succession in the presentation series would be minimal.

Out of a total of 118 sessions, 110 were given to *Ss* on separate days. Of the eight remaining sessions, a maximum of two per day for any *S* was given, with a minimum interval 4½ hr. between any two sessions.

Subjects

Eleven men and one woman participated in this experiment. They were graduate students and faculty at the University of Rochester. Two *Ss* (RH and JL) were not tested with stimulus X_1.

[3] After the first nine sessions were completed for each *S* (three controls and six experimental stimuli), it was decided to introduce stimulus X_1, in an effort to determine more closely the region of maximal after-effect. Ten *Ss* were tested with this stimulus.

Results and Discussion[4]

Controls.—The differences shown in Table I were not significant beyond the .24 level in any of the control comparisons cited in *Controls*, above, when the formula for correlated variables was used. Therefore the controls X_X and X_C were considered equivalent, and immediate and delayed test-retest reliabilities were considered sufficient.

On the basis of the findings in these control comparisons it was concluded that a combination of all three controls would lead to a more reliable measure of a true control value for these *Ss* than any one or two of the separate controls. Figural after-effects could then be calculated as the difference between these averaged controls and the mean setting made under a given experimental condition. For convenience the averaged control value will be referred to as X_C.

The mean control setting, representing the position of the two dots of the *T* figure that *S* considered horizontal, was rotated .46° in a clockwise direction from the objective horizontal. This value agrees fairly well with the value found by Rogers *et al.* (**15**). Their value, .65°, (determined by the present writer from the weighted average estimate by 53 *Ss*) of the angular position of a horizontal line whose visual angle subtended at the eye, 3.23°, was of the same order as the visual angle subtended by the two dots in the present experiment, 3.58°. In view of the wide between-subjects variability found by them and the difference in tasks, it is surprising how closely the mean constant errors agree.

Adjustment time.—The time each *S* took to adjust the *T* figure to the horizontal position was recorded for each trial. The mean time for 12 *Ss*

[4] To reduce printing costs, tables of the data for individual subjects and for order of stimuli have been deposited with the American Documentation Institute. Order Document ₁3422 from American Documentation Institute, 1719 N Street N. W., Washington 6, D. C., remitting $1.00 for microfilm (images 1 inch high on standard 35 mm. motion picture film) or $1.00 for photocopies (6 × 8 inches) readable without optical aid.

TABLE I

MEAN ADJUSTMENTS (IN MM.) AND MEAN DIFFERENCES FOR CONTROL CONDITIONS

Cond.	Session						
	1*	2**	Diff. 1 vs. 2		9**	Diff. 9 vs. 1 + 2	
			Mean	σ_M		Mean	σ_M
X_C	−.905	−.555	.097	.157	−.583	.012	.047
X_X	−.368		.083	.149			
Diff.	−.537						
σ_{M_D}	.433						

* N = 6 each condition.
** N = 12.

was 9.20 sec.; the *SD* was .67 sec. The individual means ranged from 7.88 sec. to 10.58 sec.

Because of the rapid decline of figural after-effect with presentation of the *T* figure, it is important to take account of the possible variability due to different speeds of adjusting the setting knob. Hammer (**7**) found a logarithmic decay of effect with time, with substantially zero effect beyond a 100-sec. interval between *I* and *T* figures.

In the present case 96 per cent of the settings were made within 7–11 sec. after the inspection period ended. In the Hammer study the relative decline in effect (as measured by the writer from her graphs) during this period amounted to about one-tenth of the initial after-effect. This means that the variability of the setting times in the present study was quite small in terms of the total effect, assuming a like curve of decline of effect under the present conditions. The average amount of effect at the end of 9.2 sec., the mean setting-time in this experiment, was about 75 per cent of the maximal effect, if the same assumption is made. In the present case, with different stimuli, there may be a smaller decline, since Köhler and Wallach found little difference between the amount of effect as tested with the method of constant stimuli and the method of average error (**12**).

We are interested in relative amount of effect for different stimuli, other conditions being the same. Unless a great discrepancy exists between decline rates of figural after-effect for the different stimuli, the observed mean values are probably close to the true mean values.

Cumulative effects.—The necessity for avoidance of growth of figural after-effects with successive trials by an individual has been stressed in the literature (**8, 9, 10, 11, 12, 14**). In preliminary work inder the present experimental conditions such cumulative effects were apparently present in some *S*s and not in others. A more systematic investigation was therefore made in which the first five trials of experimental sessions were compared with the last five. Table II shows the stimuli used and the mean differences and *t*'s obtained. None of the differences approaches statistical significance. Three *S*s tended to show cumulative effects; two *S*s showed a tendency toward the reverse; seven *S*s showed no marked tendency in either direction.

There are several factors which may influence the degree of cumulative effect, and these may have had their influences balanced to produce no cumulative effect in the present experiment.

TABLE II

MEAN DIFFERENCES BETWEEN FIRST AND SECOND FIVE TRIALS OF EXPERIMENTAL SESSIONS (MM.)
(N = 12)

	Stimulus					
	X_{-1}	X_0	X_1	X_3	X_4	X_5
Mean diff.	−.087	.090	.087	−.013	−.133	−.152
σ_{M_D}	.106	.115	.162	.111	.090	.129

Of these, as far as our present knowledge goes, three are of major importance. First, individuals may differ markedly in their susceptibility to accumulation of effect (**3, 11, 13, 14**).

Second, length of inspection period is of importance. Gibson and Radner (**6**) have shown that the amount of figural after-effect, in the case of linear *I* and *T* figures, increases up to about 40 sec. inspection time and then levels off. In the preliminary experimentation of this study the same results were found, leading to a choice of 1½ min. inspection time. It is reasonable to suppose, if inspection of more than 60 sec. leads to negligible increase of effect over that present with a 60 sec. inspection, that repeated inspections, each of the same duration, will likewise not result in appreciable change in amount of figural after-effect. If there existed a cumulative effect for some *S*s, it might have reached a maximum after one or two trials with 1½ min. inspection, and not appear in a comparison of the first and second five trials of any session.

A third factor is that of illumination. Freeburne and Hamilton (**3**) found that with high brightness and 5-min. inspection periods, figural after-effects tended to decline on successive days, whereas it tended to rise with low brightness. This might be an example of Köhler and Fishback's (**11**) spread of satiation to areas beyond the *T* figure—tending to reduce effects with continued presentation.

In any case, the combination of circumstances in the present experiment led to no cumulative effect, on the average. The course of cumulative effects under various conditions seems to be a fertile area for study.

The experimental variable.—The differences between the settings for each of the experimental situations and those of the combined controls are shown in Table III. The means are the average of ten trials per session for each of 12 *S*s. With the formula for correlated variables the figural after-effects for stimuli X_4 and X_5 were different from zero at the .03 and .02 levels of significance, respectively. All other differences were significant at better than the .01 level of confidence. A graphic representation of the effects is given in Fig. 2.

TABLE III

MEAN FIGURAL AFTER-EFFECTS (MM.) FOR VARIOUS DISPLACEMENTS OF *I* FIGURES FROM THE CONTROL POSITION X_C IN EXP. I

	Stimulus						
	X_{-1}	X_0	X_1	X_2	X_3	X_4	X_5
Mean	.548	.924	1.338	1.326	.638	.381	.373
σ_{M_D}	.096	.176	.190	.127	.158	.150	.140

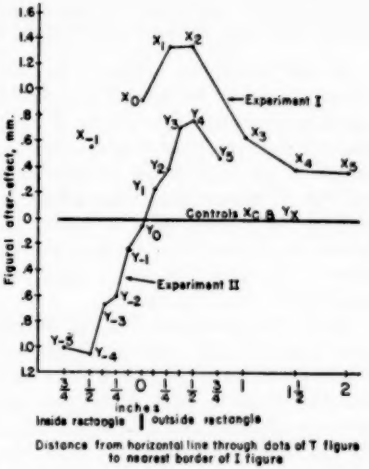
Figural after-effect, mm.
Experiment I
Controls X_C & Y_X
Experiment II
inches
Inside rectangle | outside rectangle
Distance from horizontal line through dots of T figure to nearest border of I figure

FIG. 2. Figural after-effects for various distances between *I* and *T* figures when the figure was not symmetrical with respect to the fixation cross (Exp. I), and when the *I* figure was symmetrical with respect to the fixation cross (Exp. II).

The individual changes in figural after-effect with changes of stimulus are generally of the kind represented by the mean changes. However, the absolute measure of figural after-effect can be influenced by many factors. The individual measure for any single stimulus has been found to vary with fatigue, time of day, attitude of *S* with respect to techniques of setting the dots, etc. For this reason a number of *S*s was used. No high precision is claimed for the measures as absolute ones, but on the basis of preliminary investigation the standard errors found are considered representative. Even though single sessions showed discrepancies from points on a smooth curve, repetition with a given *S* on several occasions indicates that a smooth curve of the kind described by the mean

data does result with repeated trials, although possibly with different maxima for different *Ss*.

It will be seen by reference to point X_5 on the curve for Exp. I that for the most extreme position of the *T* figure the after-effect was small, about .4 mm. With decreasing displacement of the *T* figure the after-effect increased to a maximum at point X_2, where the mean dot-displacement was about 1.4 mm. It then again decreased for X_1, X_0, and X_{-1}.

The "distance paradox" of Köhler and Wallach (**12**) states that figural after-effect increases with increased separation of *I* and *T* figures, and declines after reaching a maximum at an intermediate position. The "distance paradox" is apparently verified by these results, but only in very general terms. First, the distances tested were not close together, so that a clear picture of the course of the increase in effect to the maximum is not presented. It will be more clearly shown for satiation alone in Exp. II, by the use of smaller increments of the experimental variable. The verification of the "distance paradox" does not involve stimulus X_{-1}, since here the elements of the *T* figure lie *inside* the areas previously occupied by the rectangles (see Fig. 1). Second, an apparent contradiction to the Köhler and Wallach finding is that the figural after-effect for X_0, instead of being zero for a figure where the *T* figure lies on a border, is in the same direction as that for the other experimental stimuli. This fact and the finding for X_{-1} cast doubt on the supposition that the results of Exp. I represent the effects of satiation alone.

Köhler and Wallach's "satiation theory" would predict that a *T* figure falling on the former position of an *I*-figure boundary, other things being equal, would suffer no displacement. In X_O this is the case, except for the possible satiation influence of the upper border of the rectangle.[5] There was actually found for X_O a mean displacement of .924 mm. If we assume that the upper border of the rectangle affects the *T* figure, then comparison of the effects at this distance, $1\frac{5}{8}$ in., can be made with the effects of X_4, where the lower border is $1\frac{1}{2}$ in. away from the center of the dot in the *T* figure. In the latter case the displacement was .381 mm. The difference between the two, .543 mm., is significant at the .005 level, for 11 *df*. If we assume that the figural after-effects of light figures on a dark ground are of the same magnitude as those for dark figures on a light ground (**12**), then the results indicate either the presence of another factor or a basic inconsistency between theory and experimental results.

In the case of X_{-1} the *T*-figure elements lie inside the inspected rectangles, but closer to the border nearest the horizontal line through the fixation point. Apparent displacement took place in the same direction as that of the other experimental stimuli. Here both borders act on the *T* figure. If we make the same assumption for X_{-1} with respect to light-dark conditions, then the figural after-effect, .548 mm., should be about the same as the difference between the effects of X_2 and X_3, whose nearer borders were, respectively, $\frac{1}{4}$ in. and 1 in. from the *T* figure. This difference, .688 mm., is of the same order as .548 mm. for X_{-1}, *but in an opposite direction from that predicted.* The predicted direction of apparent displacement would be counterclockwise, whereas the apparent displacement found was actually clockwise with respect to the control setting.

It has been shown by Prentice and Beardslee (**14**), Gibson (**4, 5**), and Gibson and Radner (**6**) that a line tilted 10° from the vertical tends to adapt toward a norm (in this case the vertical) during an inspection period. That this tendency is not restricted to lines has been shown by Wertheimer (**16**) and Witkin and Asch (**17**), among others.

Many of the figures used by Köhler and Wallach actually constitute tilted figures from which one would expect adaptation in the sense of Gibson. Reference to Fig. 1 will show that all of our experimental stimuli can also be regarded as figures tilted off the horizontal. Among them X_{-1} has the least tilt and X_5 has the most. The tilt in the latter case is of the order of 45°. Based on Gibson's data (**4**) one would expect X_5 to show the least figural after-effect due to adaptation, and X_{-1} or X_0 (probably) the most. X_C would, of course, be expected to show none.

[5] In this discussion, for convenience of expression, reference is made to the right-hand rectangle of the two in the *I* figure.

The action of such adaptation takes place in a clockwise direction in the present experiment. The after-effects of this adaptation cause the dots to appear tilted in a clockwise direction after inspection, when they are actually horizontal. In order to make them to appear horizontal, *S* displaces them in a counterclockwise direction.

If it is postulated that in addition to the action of satiation, adaptation took place toward the norm of symmetry, in this case toward the position of X_C, then the results are easily explainable. The after-effects of such adaptation would lead to displacement in the direction found. To test this hypothesis Experiment II was performed.

Experiment II

The problem was to devise a set of stimuli which would test the effects of some of the distance relations between *I* and *T* figures of Exp. I more precisely, but which would be symmetrical with respect to the fixation point. Such symmetry would eliminate figural after-effects due to adaptation. The only figural after-effects, then, would be those resulting from satiation.

Apparatus and Procedure

Except for the stimuli the same apparatus and procedure were used in Exp. II as in Exp. I. The control was necessarily a stimulus consisting only of the fixation cross. The 11 experimental stimuli are shown in Fig. 3. The *I* figure in each of these consisted of the rectangle and fixation cross. The *T* figure consisted of four dots, the right-hand members of which were stationary, fixing the vertical distance. The left-hand dots were movable vertically, and were displaced equal distances, but in opposite directions during any setting. Both *I* and *T* figures were solid black. In Fig. 3, the right-hand dots were omitted for convenience in all stimuli except Y_0. The *S* was required to set the left-hand dots to a position in which all four dots formed a perfect rectangle. It will be noted that the successive distances between *I*-figure boundaries and the *T* figure were much closer together than in Exp. I, thus permitting a more precise judgment of the course of increase of effect for satiation alone.

The order of stimulus presentation was as follows: Control, Y_{-4}, Y_0, Y_{-3}, Control, Y_3, Y_4, Y_0, Y_1, Y_{-2}, Y_{-1}, Y_4, Y_{-3}, except for subject JW to whom the stimuli were presented in the order: Control, Y_{-4}, Y_0, Y_{-3}, Control, Y_3, Y_4, Y_1, Y_2, Y_{-1}, Y_0, Y_{-4}.

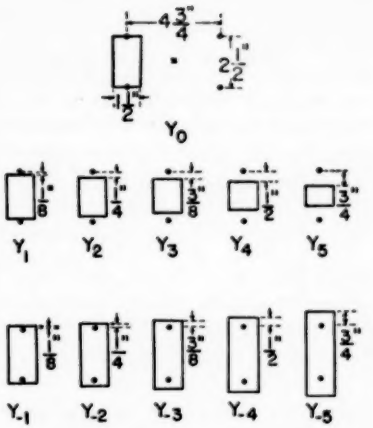

Fig. 3. Stimulus figures used in Exp. II. The fixation cross and all four dots are shown only in Y_0. Other illustrations show only the *I* figure and the left-hand dots of the *T* figure.

Six *S*s were used, two women and four men. Three of these *S*s participated in Exp. I.

Results and Discussion

The controls differed by a mean of .11 mm., with a standard error of the mean difference of .121 mm. This difference is not significant.

The figural after-effects in Exp. I and II may be compared in Fig. 2. Table IV shows the mean effects and their standard errors. The standard errors in this experiment are larger than those in Exp. I. Added to the above mentioned factors leading to variability of response, here there is also the difficulty that the adjustment by *S* must be made in terms of four dots, not two. Whatever biases toward increase of variability existed in Exp. I must now be increased to take in this additional difficulty, leading to greater standard errors. Nonetheless, the pattern of the curve for Exp. II, large negative values passing through zero to large positive

TABLE IV

MEAN FIGURAL AFTER-EFFECTS IN MM. FOR VARIOUS DISTANCES BETWEEN *I* AND *T* FIGURES WHEN THE *I* FIGURE IS SYMMETRICAL WITH RESPECT TO THE FIXATION CROSS IN EXP. II

	Distance (Inches) from Horizontal Line through Dots of *T* Figure to Nearest Border of *I* Figure										
	Outside Rectangle					0	Inside Rectangle				
	¾	½	¼	⅛	⅛		⅛	¼	⅜	½	¾
Stimulus	Y_{-5}	Y_{-4}	Y_{-3}	Y_{-2}	Y_{-1}	Y_0	Y_1	Y_2	Y_3	Y_4	Y_5
Mean	−1.00	−1.05	−.67	−.60	−.21	−.04	.25	.39	.70	.77	.50
σ_M	.251	.322	.276	.144	.139	.100	.180	.195	.127	.113	.120

values, is essentially duplicated in each *S*'s data, even though the results of single sessions may deviate from a smooth curve. It is recommended that work of this kind be done certainly with no less than six *S*s, and preferably no less than ten.

The "distance paradox" of Köhler and Wallach is supported, both for stimuli where the *T*-figure dots are inside the rectangle and outside. It is seen that the point of maximum effect is not quite the same in both curves. If such a "paradox" holds for figural after-effects resulting from adaptation (indications from Gibson and Radner's work (**6**) are that it does) then its maximum is as yet undetermined, and may have affected the results of Exp. I in an unknown way.

The hypothesis that adaptation to the norm was acting in the first experiment is supported. Tables III and IV and Fig. 2 show that the difference between the effects in Exp. I and II is consistently in the same direction.

Expecially noteworthy is the fact that the stimulus Y_0, with the dots on the rectangle border, gave effects not significantly different from zero, thus supporting that portion of Köhler and Wallach's theory.

There are at least three important influences here that might affect strength and direction of figural after-effect. The figural after-effects in the cases of X_{-1} and X_0 are different from their counterparts in the second experiment partly because of the effect of the outer border of the rectangles. Secondly, there are different tasks set in the two experiments. Gibson's experimentation (**4**) shows that figural after-effects are greater with inspections farther into the peripheral field than near the fovea. This would have led to greater effects in Exp. II than in Exp. I, instead of smaller, as found. This finding points to the third influence as being of major importance—the interaction between perceptual tendencies arising from different causes.

In the same sense that Gibson's lines (**4**) and Prentice and Beardslee's lines (**16**) adapted toward the symmetrical norm, thus producing figural after-effects on later presented lines, those in Exp. I of the present study also adapted toward positions of symmetry. It is therefore suggested that the non-symmetrical stimuli of Köhler and Wallach (**12**, Fig. 36) produced similar adaptation. This adaptation produced figural after-effects on the succeeding *T* figures, in addition to those produced by satiation. This accounts very well for the fact that there was a difference between the results of Exp. I and II, where in the latter case the effects of adaptation to the norm were eliminated.

Without claiming precision of any sort, differences might be taken between the effects of X_1 and Y_3, and between X_2 and Y_4, these being cases with equal distances between *I* and *T* figures. It is suggested that the order of magnitude of the figural after-effects of adaptation is indicated by the differences at those distances. The consequence of this reasoning is a hypothesis that Gibson's adaptation phenomenon shows

a "distance paradox" similar to that of satiation. In this case, however, the relation may be in angular terms, distance terms, or some function of the two. Interaction with satiation phenomena is present, as seen in Fig. 2, but it does not necessarily have to be of an algebraically additive nature. Further investigation is needed in this area.

Summary and Conclusions

In two experiments figural after-effects with different distances between non-symmetrical inspection (*I*) figures and a symmetrical test (*T*) figure were investigated, following one of Köhler and Wallach's figures. *I* figures, mounted on glass, were inspected against a bright background. A similarly mounted movable *T* figure was placed at different distances from the former position of the *I* figure with respect to a fixation cross, and adjusted to the apparent horizontal by *S*.

The repetition of controls showed that for present conditions controls are reliable in the same individual, but for both controls and experimental conditions great individual differences exist. The average controls of 12 *S*s showed no significant change after six sessions with experimental stimuli.

A comparison of the first five and the second five of ten trials per session showed that the cumulation of figural after-effects is not universal for some experimental conditions. Under the present conditions a decline in effect occurred as often as an increase. It is suggested that some important variables involved are stimulus brightness, duration of inspection, activity intervening between trials, and individual differences.

Since in the first experiment the results for two *I*-figure positions contradicted the "satiation theory," Exp. II was done to test the hypothesis that Gibson's adaptation to the norm was also present in the non-symmetrical situation of the first experiment, leading to interacting effects of satiation and adaptation to the norm. In the second experiment adaptation to the norm was removed by using a stimulus symmetrically placed with respect to the fixation mark. In addition, smaller increments of distance between *I* and *T* figures were used.

The "distance paradox" was generally verified in Exp. I and more precisely in Exp. II, where satiation alone was acting. Indications are that the maximum figural after-effect (satiation) for present conditions occurs with separation between *I*- and *T*-figure *boundaries* of between $\frac{1}{4}$ and $\frac{3}{8}$ in. The results in Exp. II showed consistently smaller figural after-effects than in Exp. I, thus supporting the hypothesis that both phenomena combined in some way in Exp. I. The effects of adaptation to a norm seem in the present case to be as powerful as those of satiation.

(Manuscript received
December 14, 1950)

References

1. Bales, J. F., and Follansbee, G. L. The after-effect of the perception of curved lines. *J. exp. Psychol.*, 1935, **18**, 499–503.
2. Fox, B. H. Figural after-effects: The influence of certain variables on the Köhler effect and adaptation. Unpublished Ph.D. thesis, Univ. Rochester, 1949.
3. Freeburne, C. M., and Hamilton, C. E. The effect of brightness on figural after-effect. *Amer. J. Psychol.*, 1949, **62**, 567–569.
4. Gibson, J. J. Adaptation, after-effect, and contrast in the perception of tilted lines. *J. exp. Psychol.*, 1933, **16**, 1–31.
5. Gibson, J. J. Adaptation, after-effect, and contrast in the perception of tilted lines: II. Simultaneous contrast and the areal restriction of the after-effect. *J. exp. Psychol.*, 1937, **20**, 553–569.
6. Gibson, J. J., and Radner, Minnie. Adaptation, after-effect, and contrast in the perception of tilted lines: I. Quanti-

tative studies. *J. exp. Psychol.*, 1937, **20**, 453–467.

7. HAMMER, ELAINE. Temporal factors in figural after-effects. *Amer. J. Psychol.*, 1949, **62**, 337–354.
8. KÖHLER, W. *Dynamics in psychology.* New York: Liveright, 1940.
9. KÖHLER, W., AND DINNERSTEIN, DOROTHY. Figural after-effects in kinesthesis. *Miscellanea Psychologica Albert Michotte.* Louvain: Institut Superieur de Philosophie, 1947, 196–220.
10. KÖHLER, W., AND EMERY, D. A. Figural after-effects in the third dimension of visual space. *Amer. J. Psychol.*, 1947, **60**, 159–201.
11. KÖHLER, W., AND FISHBACK, JULIA. The destruction of the Müller-Lyer illusion in repeated trials: II. Satiation patterns and memory traces. *J. exp. Psychol.*, 1950, **40**, 398–410.
12. KÖHLER, W., AND WALLACH, H. Figural after-effects: An investigation of visual processes. *Proc. Amer. Philos. Soc.*, 1944, **88**, 269–357.
13. MARKS, M. R. A further investigation into the Köhler effect. *Amer. J. Psychol.*, 1949, **62**, 62–74.
14. PRENTICE, W. C. H., AND BEARDSLEE, D. C. Visual "normalization" near the vertical and horizontal. *J. exp. Psychol.*, 1950, **40**, 355–364.
15. ROGERS, S., VOLKMANN, J., REESE, T. W., AND KAUFMAN, E. L. Accuracy and variability of direct estimates of bearing from large display screens. Special Devices Center, ONR, Memorandum Report 166-I-MHCl, May 15, 1947.
16. WERTHEIMER, M. Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.*, 1912, **61**, 161–265.
17. WITKIN, H. A., AND ASCH, S. E. Studies in space orientation: IV. Further experiments on perception of the upright with displaced visual fields. *J. exp. Psychol.*, 1948, **38**, 762–783.

CENTRAL AND PERIPHERAL FACTORS IN THE PHI PHENOMENON [1]

BY CAROL H. AMMONS
University of Louisville

AND

JOSEPH WEITZ
Life Insurance Agency Management Association

Phi movement was extensively described in 1912 by Wertheimer (**13**) who characterized it as "an unanalyzable sensation." In "explaining" this observation, Wertheimer postulated that the brain is a field in which each perceived object is represented by a physiological process (**8**, p. 54), and that this physiological correlate has a definite locus, from which circular waves of excitation diffuse outward. Thus, one process is capable of influencing other similar processes within the brain field. The spreading excitations from the two stationary stimuli, in the case of phi, occur in such a way that the experience of uniform movement is bound to them in just as *ordered* and *physically immediate* a manner, as a corresponding sensation to the action of a single stimulus (**13**).

Probably as a result of Wertheimer's assigning an important role to the stimulus situation, subsequent investigators explored mainly the differential effects of such stimulus variables as duration and intensity of stimuli, distance between stimuli, duration of dark interval, and color of stimuli and background. This literature has been critically summarized by Fernberger (**4**), Helson (**6**), Hovland (**7**), Mibai (**9**), and Neff (**10**), several of whom remark upon the lack of any thoroughly systematic study of the factors which influence the phi phenomenon. Only one generalization can really be drawn: that almost any change in the stimulating agencies may arouse the perception of motion in an attending *S*. As a rule, functional relationships have not been determined.

Some attempt has been made to determine the importance of peripheral and central factors in the perception of the phi phenomenon. A study by Gengerelli (**5**) was designed to discover whether phi is more frequently observed when both stimuli give rise to excitations in the same cerebral hemisphere than when they cause excitations in different hemispheres. Statistically reliable differences were found between the conditions of stimulations of the same and opposite hemispheres, and led Gengerelli to conclude that two excitations existing in the same hemisphere are more "strongly interrelated" than two excitations existing in different hemispheres. Although this is an excellent approach to the problem one important variable was not investigated, that is, the relative role of retinal factors in mediating the phi phenomenon.

Shipley, Kenney, and King (**12**) reported that phi movement was perceived when stimulation was monocular, binocular, and interocular. They

[1] This study was done at Newcomb College, Tulane University. The authors wish to express their appreciation to Dr. Dorothy W. Seago for her assistance throughout the study.

controlled for general binocular stimulation effects, but failed to take into consideration the possible effects of exciting one or both cerebral hemispheres by stimulating corresponding or opposite halves of the two retinas.

In exploring the relationship between illumination factors and movement perception, Bartley (**1,2**) came to the conclusion that retinal factors facilitate spatial summation and consequently are important in the perception of phi.

Problem

If one assumes along with Köhler that " . . . the neural functions and processes with which the perceptual facts are associated in each case are located in a continuous medium" (**8**, p. 55) and that the environment is homogeneous (**8**, p. 67) then the perception of phi movement would be equally frequent when peripheral stimulation was of one or both eyes, and when this stimulation gave rise to excitations in the same hemisphere or in different hemispheres of the brain. However, if retinal factors appreciably influenced the perception, differences would appear in the frequency and quality of phi movement reported under these different conditions of stimulation. The problem, then, is to investigate the relative contribution of peripheral and central factors in the perception of phi movement.

Experimental Design

Four conditions of observation were set up to test the above predictions.

Condition I.—The *S* viewed both stimuli monocularly with his dominant eye, while fixating a just visible red cross. The stimuli for phi were 10° apart, 5° out to the right and left of the central fixation point.[2]

Condition II.—The *S* fixated the red cross binocularly; however, a septum with graduated wings was placed in such a way that each eye was presented separately with a single stimulus 5° out from the fixation point. This condition served as a check against Cond. I permitting evaluation of retinal interaction as a contributing factor to the perception of phi.

Condition III.—Since both Cond. I and II resulted in stimulation of opposite cerebral hemispheres, two other conditions were necessary in order to evaluate the peripheral and central factors involved. In Cond. III the two stimuli, still separated by 10°, were presented monocularly. Both stimuli were presented on the same side of the fixation point. One was 5° from the fixation point, the other 15°.

Condition IV.—This condition involved the same physical arrangements as Cond. III except that, as in Cond. II, a septum allowed the presentation of one stimulus to each eye. In both Cond. III and IV the same hemisphere was stimulated.

Method

Subjects.—The *S*s were 22 graduate and senior psychology students, 14 men and 8 women. Sixteen reported having uncorrected 20/20 vision in both eyes. Four *S*s who reported corrected astigmatism and two *S*s with corrected myopia, wore their glasses during the experimental sessions. All *S*s were trained observers, but were naive with respect to the purpose and methods of this experiment.

Apparatus.—The *S*'s head rested comfortably in a headrest which had been adjusted to his height. On a table approximately 1 in. in front of his face was a panel with two rectangular openings through which he could look. A second panel, 19 in. from *S*'s eyes, contained small openings in which the phi stimuli appeared. Also in this panel, 1.5 in. above the level of the stimuli, was an opening for the fixation stimulus. All of the apparatus was painted flat black.

The panel containing the two rectangular holes at eye level served a dual purpose. It minimized reflected light and had a cardboard disc pivoted between the two holes which enabled *E* to block the vision of one of *S*'s eyes quickly and easily without his knowledge. This arrangement permitted *S* to keep both eyes open at all times.

The fixation stimulus appeared in the center of the second panel. It was a small red cross, cut in Hering black paper, fastened to a red

[2] Since the question of macular representation in the cerebral cortex is unsettled as yet, stimulation of the macular area was avoided by using this 10° separation, based on Polyak's report (**11**) that the combined area of the human fovea and macula lutea subtends a visual angle of 5°.

Wratten filter (No. 25). The intensity of the fixation stimulus could be varied by an optical wedge.

In the same plane, but 1.5 in. below that of the fixation stimulus, were four openings, 5° and 15° to the right and 5° and 15° to the left of the fixation point. With this arrangement the stimuli were always separated by 10°. The stimuli, two vertical slits, 6 mm. by .5 mm., were cut in Hering black paper and fastened to the ends of covered lucite tubing (.5 in. in diameter) through which the light was "piped" from metal housings. The light sources were 110-v., 7-w. bulbs. A simple cam arrangement driven by a 60 rpm Telechron synchronous motor activated two microswitches, each of which completed the circuit for one of the lights.

A septum with three wings graduated in length was placed between the two panels. It was used to cut down reflected light and to separate the vision of the two eyes for Cond. II and IV.

Procedure.—The experiment was conducted in a small windowless room painted entirely with flat black paint. A ventilating fan ran at all times and served to mask the rhythmic sound of the microswitches. Cotton wax was placed in *S*'s ears to minimize irrelevant auditory stimulation.

The 22 *S*s were observed once under each of the four experimental conditions. In order to equalize the effects of experience a random order of presentation of the sequences of conditions was used. Any two experimental sessions for a given *S* were separated by more than a 2-hr. interval, and most pairs of sessions were separated by approximately 24 hr. Sessions were 1 min. in duration. At the beginning of each session *S* was dark adapted for 10 min. At the end of this time *E* adjusted the optical wedge until *S* reported being just able to see a stable red cross. Then *E* asked *S* to describe in detail the movement which he saw. During the following 1-min. period the stimuli flashed 60 times, the flashes separated by approximately 80 msec.[3]

The *E* recorded *S*'s description and scored it. If the response could not be scored, further questions were asked until it could be. Scoring consisted of characterizing the perceptual experience as (*a*) consistent phi movement observed throughout the stimulation period, (*b*) no phi movement observed, or (*c*) phi movement observed during only part of the observation period. Phi movement was judged to have been seen when a smooth movement of any kind was reported from the initial to the final light position.

[3] This was the interval at which phi was most often reported by a preliminary group of ten *S*s.

TABLE I

FREQUENCY OF REPORTS OF PHI AND JUDGED DISTANCE OF MOVEMENT UNDER FOUR CONDITIONS OF STIMULATION

Condition	Frequency of Report			Judged Distance in Inches	
	Phi	No Phi	Both	Median	Mode
Opposite hemispheres					
Monocular	7	4	11	3.5	3
Binocular	2	10	10	3.5	3
Combined	9	14	21	3.5	3
Same hemisphere					
Monocular	11	5	6	3.0	1
Binocular	4	10	8	2.0	1
Combined	15	15	14	2.6	1
Combined monocular	18	9	17	3.2	3
Combined binocular	6	20	18	3.2	3

After each experimental session *S* was asked to estimate the distance through which the light had moved or the distance between the lights.

RESULTS

From Table I it is apparent that phi movement was perceived, at least by some *S*s, under each of the four conditions as predicted from Köhler's theory (**8**). However, all *S*s did not report perception of phi under all conditions.

It was thought desirable to obtain a rough estimate of whether these frequencies of the different kinds of report varied from condition to condition more than could be attributed to chance. An over-all chi-square of 13.60 was significant at the .05 level of confidence.[4] In other words, it would seem that some factor other than chance was operating to produce this difference in the reports.

A comparison of combined conditions of monocular stimulation with

[4] Caution should be exercised in the interpretation of this result since the underlying assumptions of chi-square are not met. Since each *S* served in each of the four conditions, the measures are related.

combined conditions of binocular stimulation as shown in Table I indicates that the frequency of report of phi was greater under monocular stimulation than under binocular stimulation (18 vs. 6). If we combine these frequencies with those of *Ss* who reported alternation of phi with other kinds of movement or no movement there is still a small difference between the two conditions, again favoring monocular stimulation.

Qualitative analysis of *Ss* reports showed that when phi was reported under binocular presentation it tended to be less stable than the phi reported under monocular conditions. This manifested itself in reports of fluctuations from perception of phi to other kinds of movement, alternate flashings of the lights, jerky movements or "successive jumps or steps" through the field.

The frequency of perception of phi alone was slightly greater when both excitations were located in the same cerebral hemisphere than when the excitations were located in opposite hemispheres (15 vs. 9).

It may be seen from Table I that when stimulation was of the same hemisphere the estimated distance of the phi movement was less than when stimulation was of opposite hemispheres, regardless of whether the peripheral stimulating conditions were monocular or binocular. However, caution must be exercised in interpreting both the median and modal values since there were three extreme values given by two *Ss* under different conditions.

It is interesting that where *S* reported seeing no movement at all under one of the conditions the estimated separation of the two stationary lights was usually quite close to the real physical distance. However, the estimated distance changed when phi was reported and was usually judged smaller than the physical separation under all conditions, although several *Ss* reported distances from 5 to 12 in. Partial movement of each stimulus was described several times as covering from 1 to 2 in., with a gap of unspecified length in between.

The descriptions of phi were varied but easily identifiable: "smooth like a pendulum swinging . . . ," "back and forth motion of a small bar . . . ," ". . . a dot of light which dips downward in an arc . . . ," ". . . one light swinging as though I am on a ship at sea . . . ," ". . . one flash of light moving across the field, but it has bumpy edges," ". . . definitely has an impression of wholeness or unity . . . ." One *S* reported ". . . it looks like a baseball bat flipping end over end. . . ."

Several kinds of movement other than phi were reported. These included partial or dual motion, movement of a flash of light between the two stimuli, as reported by Dimmick (**3**), and circular or tri-dimensional movement. Sometimes alternation of several of these during a single observation was reported. Partial motion was described as, ". . . there's movement, but I can see that there are two lights. One seems to move part way but goes out, then the other comes on; it needs to be faster and then it would move all the way across. . . ." The *Ss* described the "Dimmick" flash as, ". . . the lights are alternate, but there seems to be a wavy or foamy movement between them now." One *S* reported seeing "a filmy cloud [between the two lights] which is circular in shape . . . the flash moves but not the two lights . . . it is a little blob glowing." The circular movement was usually quite startling to *Ss*. Some reported the direction to be clockwise, others, counterclockwise. One *S* said, "The light looks like an arrow or dagger with a reddish tinge moving in a circle as though it were turning on a kymograph drum."

Several other observations of possible significance were made. Only three *Ss* were aware that some conditions were monocular. These either blinked their eyes or by accident touched the panel which blocked their vision.

All *Ss* expressed surprise when allowed to see the apparatus after their fourth and last experimental session. All but one *S* reported that they had thought that the stimuli were much farther away. It is interesting that the *S* who thought the stimuli were 8 to 10 in. away did not report phi under any condition. Two other *Ss*, who imagined the stimuli to be ten or more feet away perceived phi under all conditions.

There were no apparent sex differences in frequency or quality of the perceptions.

Discussion

Since the perception of phi movement did not occur with equal ease for all *Ss* under all of the four conditions of observation, a homogeneous brain field theory of perception probably cannot be accepted without reservations. Analysis of the results indicates that we must consider retinal factors as well as cortical factors in the perception of phi movement.

It has been shown in this study and in a study by Shipley (**12**) that phi movement can be perceived with binocular as well as monocular stimulation. Bartley (**2**) suggests that mutual facilitation possibly may occur in the optic tracts above the optic chiasma where the fibers from the two retinas are going to the same hemisphere, although there is no actual synaptic junction. It would seem that this interaction may take place in the ipsilateral geniculate bodies, but not in the contralateral geniculate bodies (cf. Cond. II) since Polyak (**11**) demonstrated that no retrograde degeneration of neurons in the contralateral geniculate body takes place as a result of lesions higher up in the optic tract. It is quite possible that interaction also occurs on a cortical level.

The fact that the frequency of phi perception is greater with monocular stimulation than with binocular and that the perceptions under the former conditions (I and III) are more stable indicates the importance of subcortical neural levels. It is probable that spatial summation, which occurs at the level of the second and third order neurons in the retina, is involved. This is definitely suggested by Bartley (**2**).

The probable influence of retinal factors in the perception of phi is further hinted at by the results from the monocular stimulation conditions. A higher frequency of phi movement was reported in Cond. III where the stimuli were farther out in the periphery than in Cond. I. This would be predicted from neural structure. It is well known that in the perifoveal region which was stimulated in Cond. I, the relationship of receptor cells to the underlying neural cells is not quite one to one. Furthermore, from 10° to 15° out peripherally, there is an increase in the number of receptor cells which converge upon a single bipolar or ganglionic cell. Hence, stimulation of the periphery, as in Cond. III, probably results in greater spatial summation effects. One would then predict the perception of phi movement to be facilitated to the extent that it is based upon spatial summation.

Spatial summation effects may also serve as the basis for the observed tendency to make shorter estimates of distance under conditions of stimulating the same hemisphere (Cond. III and IV) and also for the increased variability in these estimates. Gengerelli (**5**) has pointed out that excitations located in the same hemisphere seem phenomenally to be more closely related than those located in opposite hemispheres. These results are difficult to predict from Köhler's (**8**) field theory.

Summary Conclusions

It was the purpose of this experiment to investigate the relative role of central and peripheral factors in the perception of phi movement. Twenty-two *Ss* observed phi under each of four conditions. These conditions were designed to allow comparison of movement using monocular and binocular stimulation of the same and different cerebral hemispheres.

Phi movement was reported in all conditions of observation, but the greatest number of reports of uncomplicated phi movement occurred with monocular stimulation. Although phi was perceived with binocular stimulation, reports were more variable and there was more fluctuation and variety in kinds of motion perceived. It was also found that when stimulation was of the same cerebral hemisphere, phi tended to be judged as moving a shorter distance than when different hemispheres were stimulated, even though the time and distance between stimuli were constant.

It is concluded that retinal or at least subcortical factors are important in the perception of phi and it would seem that a homogeneous brain field theory is too general and oversimplified to account for the present findings.

(Manuscript received
December 9, 1950)

References

1. Bartley, S. H. Relation of retinal illumination to the experience of movement. *J. exp. Psychol.*, 1936, **19**, 475–485.
2. Bartley, S. H. *Vision.* New York: Van Nostrand, 1941.
3. Dimmick, F. L. An experimental study of visual movement and the phi phenomenon. *Amer. J. Psychol.*, 1920, **31**, 317–332.
4. Fernberger, S. W., *et al.* Perception. *Psychol., Bull.*, 1941, **38**, 432–468.
5. Gengerelli, J. A. Apparent movement in relation to homonymous and heteronymous stimulation of the cerebral hemispheres. *J. exp. Psychol.*, 1948, **38**, 592–599.
6. Helson, H. The psychology of Gestalt. *Amer. J. Psychol.*, 1925, **36**, 342–370, 494–526; 1926, **37**, 25–62, 189–223.
7. Hovland, C. I. Apparent movement. *Psychol. Bull.*, 1935, **32**, 755–778.
8. Köhler, W. *Dynamics in psychology.* New York: Liveright, 1940.
9. Mibai, S. An experimental study of apparent movement. *Psychol. Monogr.*, 1931, **42**, whole No. 190.
10. Neff, W. S. Critical investigation of visual apprehension of movement. *Amer. J. Psychol.*, 1936, **48**, 1–42.
11. Polyak, S. *The retina.* Berkeley: Univ. Calif. Press, 1941.
12. Shipley, W. C., Kenney, F. A., and King, M. E. Beta apparent movement under monocular, binocular, and interocular stimulation. *Amer. J. Psychol.*, 1945, **58**, 545–549.
13. Wertheimer, M. Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.*, 1912, **61**, 161–278.

THE EFFECT OF REINFORCEMENT ON CLOSELY FOLLOWING S–R CONNECTIONS: II. EFFECT OF FOOD REWARD IMMEDIATELY PRECEDING PERFORMANCE OF AN INSTRUMENTAL CONDITIONED RESPONSE ON EXTINCTION OF THAT RESPONSE [1]

BY MOHAMED O. NAGATY

Yale University

The present experiment is the second of two studies designed to determine the effect of reinforcement on closely following stimulus-response (S-R) connections. The first paper (**13**) discussed the controversy among psychologists concerning the effect of reinforcement on following S-R connections either in "spread of effect" or in "backward conditioning" experiments, as well as the results from Pavlov's laboratory indicating that the backward conditioning procedure produces extinction. The experiment described in the first paper showed that electric shock as a reinforcement had no appreciable effect on an immediately following conditioned avoidance response, namely, wheel rotation to the sound of a buzzer. The present experiment studies the effect of food as a reinforcement on an immediately following instrumental conditioned response (CR), namely, bar-pressing.

The conflicting results of "spread of effect" experiments with human *Ss* (**2, 9, 16, 17, 20, 21, 22, 24**) are probably due to the difficulty of controlling all the disturbing factors (e.g., rehearsal and guessing habits), a fact which makes the direct measurement of the effect of reinforcement extremely difficult. On the other hand, the results of "spread of effect" experiments using mazes with rats (**8, 12, 19**) are also subject to doubt because other investigators (**1, 3, 4, 15, 18, 23**) give ample evidence of the existence of many disturbing factors (e.g., anticipatory and perseverative responses) in maze experiments.[2]

The writer believes that in the present experiment where the direct effect of a food reward upon the immediately following bar-pressing response is studied, many of the disdurbing factors found in other experiments are avoided.

[1] One of two studies submitted to the faculty of the Department of Psychology of Yale University in partial fulfillment of the requirements for the Ph.D. degree. The writer is greatly indebted to Fouad First University of Cairo, Egypt for financial support; and to Professor Neal E. Miller, under whose guidance this investigation was performed, for his generous advice and valuable suggestions throughout the course of the experiment and the preparation of the manuscript. The writer wishes also to thank Dr. Harry G. Yamaguchi for his assistance in constructing the experimental apparatus.

[2] It should be noted that in one of the maze experiments with rats by Jenkins (**8**), where positive results of "spread of effect" were reported, anticipatory and perseverative responses were also found to yield results similar to the curve of repetitions produced by the "spread of effect." In another maze experiment by Muenzinger, Dove, and Bernstone (**12**), where positive results of "spread of effect" were also reported, the *Es* also admitted the existence of some disturbing factors producing the repetition of some responses.

Subjects And Apparatus

Subjects.—The *Ss* were 41 male albino rats 60 to 90 days old and without previous experimental training at the beginning of the experiment. Four additional rats were discarded for failure to learn the bar-pressing response during the preliminary training. The *Ss* were randomly assigned to experimental and control groups.

Apparatus.—The apparatus was a modified Skinner box. It was carefully soundproofed, lighted with a 6-w. electric bulb and provided with a glass observation window at the top. It contained two inner compartments separated by a brass panel; one compartment for the animal, and the other for the bar and the food release mechanisms. The animal compartment was made as small as possible ($6 \times 5 \times 4\frac{1}{2}$ in.) in order to reduce the random behavior of *Ss* to a minimum. The brass panel had an aperture through which a horizontal, cylindrical bar could be inserted and retracted automatically by means of two magnets mounted behind the panel. When the bar was inserted it extended into the animal compartment 1 in. beyond the panel and $3\frac{7}{16}$ in. above the food dish which was mounted on the floor. A slight downward pressure on the bar closed a microswitch which could be used to actuate the food-delivery mechanism. This released a pellet of wet food (used because the animals could consume it more quickly than dry food) and produced a moderately loud click. The presentation of the bar and the food was controlled automatically and could be adjusted so that either the food or the bar was presented first. The latencies of the bar-pressing responses were measured by a Springfield electric clock, and the number of responses was recorded by an electric counter. The length of the intervals between the presentation of the bar and the delivery of food was controlled by an electronic timer (7).

Procedure

Habituation.—The *Ss* were fed for four days at the same hour at which the experiment later took place. On Days 1 and 2 they were fed in their living cages; on Days 3 and 4 they were fed in habituation boxes made exactly similar to the experimental box. During these four days of habituation the *Ss* were fed dry food made of Purina dog chow, and no water was available during the feeding hour.

Preliminary training.—In order to prepare them for the experiment proper, all *Ss* were trained to press the bar immediately after its insertion into the apparatus. This preliminary training was accomplished in a number of stages.

On Day 5 the *Ss* were trained to get pellets of wet food (weighing approximately .12 gm. each) from the food release mechanism. The bar was not present during this training. At first they were allowed to eat ten pellets of food that had already been placed in the food dish. While they were eating, the empty food release mechanism was operated every 5 sec. in order to habituate them to the loud clicks that it produced. After this they received 80 pellets, each of which was released by the food release mechanism. During this training a new pellet was released as soon as the old one was consumed, but after every 20 pellets the *Ss* were removed to their living cages for a 3-min. rest.

On Day 6 each *S* was given five more pellets by the food release mechanism. After this the bar was inserted (smeared with a little wet mash the first time in order to attract attention) and *S* had to press it to secure a pellet. Each *S* was left in the apparatus until 80 pellets had been obtained by pressing the bar. After each 20 bar-pressing responses, *E* retracted the bar for a 2-min. interval.

During the first part of the session on Day 7, the rat was put in the apparatus with the bar already present. The fifth pressing response caused the bar to retract automatically. From this point on the bar was introduced by *E* at intervals of 20 sec. and the pressing response by *S* caused it to retract at once automatically. Twenty more pellets were given in this way. Then the *Ss* were divided into two groups.

Two preliminary training groups.—The purpose of these two groups was to control for the possible effect of a generalization decrement produced by differences between the conditions of preliminary training and experimental testing. Since the two most important conditions of experimental testing were to be (a) the food pellet given 2 sec. before the bar was presented, and (b) the food pellet given 20 sec. before the insertion of the bar, half of the *Ss* (randomly selected) were given the remainder of their preliminary training under each of these two conditions. In both groups a pellet was also given immediately after pressing the bar. Thus the 2-sec. group received one pellet immediately before and another immediately after pressing the bar.

The interval of 2 sec. before the presentation of the bar was selected to give the most immediate association possible between the reward and the subsequent pressing of the bar since the *Ss* required approximately 2 sec. to consume the food. The interval of 20 sec. before presentation of the bar was selected to give little or no association between the reward and the subsequent pressing of the bar. During this part of the training, trials were given to both groups at 50-sec. intervals. Thus the additional reward given 20 sec. before the next response occurred

approximately 30 sec. after the last one; the additional reward given 2 sec. before the next response occurred approximately 48 sec. after the last one.

During the second part of Day 7, both groups were given 20 trials with the appropriate procedure. The latencies of the pressing responses were measured by a Springfield electric clock.

On Day 8 the *Ss* of both groups were given 40 more trials with the appropriate procedure. By the end of Day 8 all *Ss* had learned to press the bar as soon as it was presented.

Experimental procedures.—On Day 9 the main experiment started. Each of the two preliminary-training groups was divided randomly into the following four experimental groups:

1. In order to test for the effects of a food reward given immediately before a response, *Ss* of Group 1 were given a pellet 2 sec. before the insertion of the bar and were given no pellets after pressing it.

2. In order to have a control tested under exactly the same conditions but without a close association between the reward and the subsequent response, *Ss* of Group 2 were given a pellet 20 sec. before the insertion of the bar and were given no pellets after pressing it.

3. In order to determine what the general effect of a food reward (administered either immediately before the performance of the bar-pressing response or between trials) was on *Ss* of the two preceding groups, *Ss* of Group 3 were not given any pellets at all either before or after pressing the bar. This is, of course, the normal procedure of experimental extinction.

4. In order to demonstrate that the time between getting the food and pressing the bar in the first experimental group was not too long for reward to have an effect when administered *after* the response, *Ss* of Group 4 were given a delayed reward after pressing the bar and were not given any pellets before pressing it. In order to be sure that the delay involved in this group was at least as long as the interval between reward and subsequent bar pressing in the Group 1, each *S* in this group was matched with a partner in Group 1. During the first ten experimental trials the length of the delay in the reward of each *S* in Group 4 was determined by the time between the delivery of food and the pressing of the bar by his partner. It should be noted that for *Ss* in Group 1 this time included the fixed 2-sec. interval, most of which was spent eating the reward, as well as the time taken to press the bar after its insertion. If it is assumed that all of the reward occurred at the moment of its delivery, then the intervals between reward and response would be equal in both groups; if it is assumed that some of the reward occurred while the animal was eating, then the intervals between reward and response would be *shorter* for Group 1 than for Group 4. Since exploratory work indicated that *Ss* in Group 1 were going to slow down and eventually stop responding, the latencies beyond the first ten trials were not used for matching; instead, the latencies on the first ten trials were repeated again during successive blocks of ten trials.

On Days 9 and 10 the *Ss* of each group were given 50 experimental trials spaced at 50-sec. intervals. During all these trials the insertion of the bar and the release of pellets were controlled automatically. When bar-pressing was not to be rewarded, the bar was disconnected from the food release mechanism so that no secondary reinforcement from the sound of its click was possible. Arrangement was also made so that the bar would retract immediately when it was pressed; but if *S* failed to press the bar it would stay exposed for a period of 10 sec. and then would retract automatically. Latencies and frequencies of pressing the bar were recorded. If *S* did not press the bar during the

TABLE I

SUMMARY OF EXPERIMENTAL DESIGN

Experimental Procedure	Preliminary Training Procedure: Food Reward Immediately before CR and Immediately after CR	Preliminary Training Procedure: Food Reward 20 Sec. before CR and Immediately after CR	*N*
1. Food reward immediately before the CR	6*	5	11
2. Food reward 20 sec. before the CR	6	5	11
3. No reward (normal extinction)	6	5	11
4. Food reward at various intervals after the CR**	4	4	8
N	22	19	41

* Number of rats in each group.
** Matched to time between food and bar-pressing of partner in Group 1.

10-sec. period when it was present, an arbitrary latency score of 10 sec. was given.[3]

Summary of design.—The general design of preliminary training and experimental trials is summarized in Table I.

Test trials.—In using the performance during the 100 experimental trials to compare the different groups, it is impossible to separate the immediate distracting or facilitating effects of the procedure (e.g., a pellet of food immediately before the insertion of the bar) upon the performance of the habit during these trials from any more permanent learned after-effect of reinforcement or inhibition. In order to separate the effects on learning from those on immediate performance, it was necessary to test *S*s from each group under the same conditions. This was done on Day 11. The *E* put the *S*s from each group in the apparatus, and after 2 min. inserted the bar and kept it exposed to the *S*s for a period of 20 min. The number of bar-pressing responses was recorded every minute. During all the test period the food magazine was empty, and the bar was disconnected from the food release mechanism so that no primary or secondary reinforcement was possible.

Results

Because the distributions of the raw latency scores were skewed, it was thought better to transform those scores into their common logarithm before any statistical test was made. This transformation (11) has the effect of normalizing the distributions of latency scores and equalizing the variances so that the familiar statistical tests become possible.[4]

Preliminary training.—In order to determine the effects of the preliminary training on the performance under the experimental conditions, the scores (for log latency and frequency respectively) that each *S* made during his 100 experimental trials were summed and an analysis of variance was made of the effects of the two preliminary training and first three experimental conditions.[5] The results are shown in Table II. As can be seen, no significant differences were found among the after-effects of the two preliminary training procedures on either log latency or frequency scores. Furthermore, there was no reliable interaction. Additional analyses showed that there was no reliable difference between the effects of those two preliminary training procedures upon the performance of the different groups during either the 40 trials of the last day of preliminary training ($p > .50$), or during test trials (see Table III). Thus, receiving a reward during preliminary training before (in addition to after) pressing the bar does not seem to have produced any stronger response. Because of this lack of difference, the groups that received identical experimental procedures but different preliminary training procedures were combined in all subsequent analyses of the data.

Experimental trials.—Figure 1 presents the mean log latencies of CR's during the 100 experimental trials for each of the four experimental groups; Fig. 2 presents the mean frequencies. From these figures it is clear that the curves for the non-reinforced group and for the two groups with reinforcement either immediately or 20 sec. before the presentation of the bar showed experimental extinction: the latency of CR's increased and their frequency decreased. As can be seen from Table II, none of the differences

[3] Although there is some indeterminateness in this arbitrary score, it will be very slight when the raw low latency scores are transformed into their common logarithm.

[4] Furthermore, for each *t*-test or *F*-test made, chi-square tests were also made, and in all cases there was almost complete agreement among the results of those tests.

[5] The control group, with a delayed reward during the experimental trials, was omitted from this analysis because the performance of these animals (showing no extinction) was so different from that of those in the other three groups. No differential effects of the two conditions of preliminary training were apparent in this control group.

TABLE II

ANALYSIS OF VARIANCE OF PERFORMANCE DURING THE 100 EXPERIMENTAL TRIALS

Sources of Variation	*df*	Log Latency Scores		Frequency Scores		$F_{.05}$
		Variance	*F*	Variance	*F*	
Preliminary training conditions	1	.014	1.28	211.87	1.28	4.21
Experimental conditions	2	.032	2.99	430.21	2.59	3.35
Interaction	2	.002	6.50	32.22	5.15	19.46
Error	27	.011		165.94		

among these three groups (for either log latency or frequency scores) was statistically reliable.

It is clear from these results that there is little if any difference between the *S*s with reinforcement immediately before the presentation of the bar and those with non-reinforcement or with reinforcement 20 sec. before the presentation of the bar. This indicates clearly that food reinforcement does not strengthen the immediately following S-R connections as would be expected from the principles of "a bi-directional spread of effect" or "backward conditioning." On the other hand, although experimental extinction occurs under these conditions in accordance with Pavlov's report (**14**, pp. 242–244), the immediately preceding food reinforcement does not seem to hasten this extinction; it seems to be without any specific effect.

While food reinforcement did not strengthen the immediately *following* S-R connections, it strengthened greatly those connections which *preceded* it. It is clear from Fig. 1 and 2 that the latency curve of the group with the delayed reinforcement decreased and the frequency curve of that group maintained a maximum level of performance during the 100 experimental trials. The differences (for both log latency and frequency scores) between this group and its matched group with reinforcement immediately before the

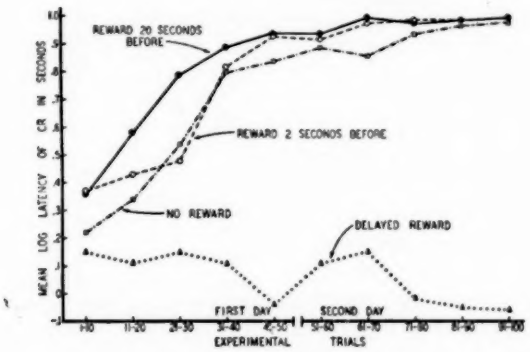

FIG. 1. Mean log latency of CR of the four experimental groups during 100 experimental trials on two successive days.

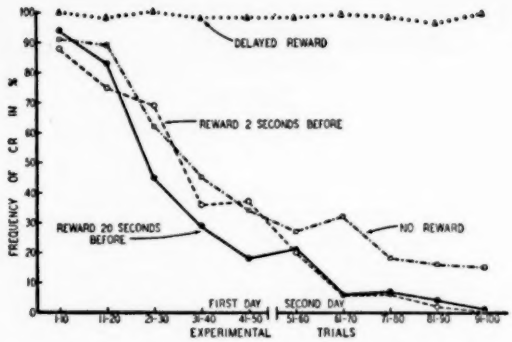

FIG. 2. Mean frequency of CR of the four experimental groups during 100 experimental trials on two successive days.

presentation of the bar were highly reliable ($p < .001$) for both log latency and frequency scores.

Test trials.—The number of CR's of the delayed reinforcement group (Group 4) was very much higher (Mean = 65.3) than those of the other three groups (Means of 19.5, 17.4, and 17.1 for Groups 1, 2, and 3, respectively) during the 20-min. test period without reward on Day 11. The difference in performance during the test period between the delayed reinforcement group (Group 4) and its matched group with reinforcement immediately before the presentation of the bar (Group 1) was highly reliable ($p < .001$). None of the differences among the other three groups was reliable, as can be seen from Table III.

TABLE III

ANALYSIS OF VARIANCE OF PERFORMANCE DURING TEST TRIALS

Sources of Variation	*df*	Variance	*F*	$F_{.05}$
Preliminary training conditions	1	0.67	208.60	249.50
Experimental conditions	2	20.36	6.86	19.46
Interaction	2	131.49	1.06	19.46
Error	27	139.76		

The results of the test trials when the after-effects of the different experimental procedures are tested under equivalent conditions, show clearly that there is little if any difference between the group which was reinforced immediately before the presentation of the bar and the other two groups either with reinforcement 20 sec. before the presentation of the bar or with non-reinforcement. It is clear from the similar performance of these three groups during test trials that food reinforcement does not have any marked effect upon the immediately following bar-pressing response. It does not strengthen it; it even does not reliably retard extinction, nor does it hasten extinction. By contrast, the delayed food reinforcement was found to strengthen greatly the remotely preceding bar-pressing response. Thus, the conclusion is that food reinforcement has little if any specific effect on immediately following S-R connections, while it strengthens the remotely preceding S-R connections.

Discussion

In the present experiment food was used as reinforcement for the learning and maintenance of the instrumental response of bar pressing. In a previous experiment (**13**) a different type of reinforcement, electric shock, was used for the learning and maintenance of an avoidance response (wheel turning) conditioned to a buzzer.[6] Exactly the same results were secured for these two different types of reinforcement. In the present experiment, when the reinforcement of food followed the response of pressing the bar to the cue of its insertion, the rats learned that response. Furthermore, a delayed reward of food after pressing the bar maintained that response. But when the direction of the association was changed by administering the food immediately before the response, it was extinguished. In the previous experiment (**13**), administering the shock after the buzzer served as a reinforcement to establish and maintain turning the wheel in response to the buzzer. When the direction of the association between the shock and the response was changed by giving the shock 1 sec. before the response of turning the wheel to the cue of the buzzer, that response was extinguished. In these experiments neither of the two different reinforcements (food or electric shock respectively) had any appreciable effect when administered before the response: they did not retard experimental extinction, nor did they hasten it.

[6] Although there is a theoretical dispute about whether the electric shock serves as a reinforcement for conditioned avoidance because it elicits the responses (**5**), or because the end of the shock constitutes a reduction in drive (**6**, **10**), there is no doubt that the shock has the empirical properties of a reinforcement in that its presence after the CS results in a strengthening of the CR and its withdrawal results in the extinction of the CR.

Summary

Forty-one albino rats were trained first to press a bar as soon as it was presented to them in order to secure a small pellet of wet food, and then were given 100 experimental trials under different conditions: (a) Group 1 was given a pellet of food 2 sec. before the presentation of the bar, and no food after pressing the bar. (b) Group 2 was given a pellet of food 20 sec. before the presentation of the bar, and no food after pressing the bar. (c) Group 3 was not given any food, either before or after the presentation of the bar. (d) Group 4 was given a pellet of food at various intervals after pressing the bar, and no food was given before its presentation. The *S*s of this group were matched with the *S*s of Group 1 with respect to the actual time interval between pressing the bar and delivery of reinforcement.

From the results of the present experiment we arrive at the following conclusions:

1. The effects of giving a food reinforcement immediately before the presentation of the bar are to produce a decrement in the strength of the bar-pressing response (confirming Pavlov's report) rather than to produce any increment as might be expected from the principles of "backward conditioning" or "a bi-directional spread of reinforcement."

2. The amount of this decrement is not reliably different from that observed in a control group which was not reinforced, or in another control group which was reinforced 20 sec. before the presentation of the bar. Thus, the direct association between food reinforcement and the following instrumental CR seems to have been without any specific effect upon that response.

3. While food reinforcement seemed to have been without any specific effect when it *preceded* the instrumental response by a brief interval, it had a marked effect when it *followed* the response after an interval of the same length.

(Manuscript received
December 6, 1950)

REFERENCES

1. BALLACHEY, E. L., AND BUEL, J. Centrifugal swing as a determiner of choice point behavior in the maze running of the white rat. *J. comp. Psychol.*, 1934, **17**, 201–223.
2. BRANDT, H. The spread of the influence of reward to bonds remote in sequence and time. *Arch. Psychol., N. Y.*, 1935, **27**, No. 180.
3. BUEL, J. Differential errors in animal mazes. *Psychol. Bull.*, 1935, **32**, 67–99.
4. DASHIELL, J. F. Directional orientation in maze running by the white rat. *Comp. Psychol. Monogr.*, 1930, **7**, No. 2.
5. GUTHRIE, E. R. *The psychology of learning.* New York: Harper, 1935.
6. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
7. HUNTER, T. A., AND BROWN, J. S. A decade-type electronic interval timer. *Amer. J. Psychol.*, 1949, **62**, 570–575.
8. JENKINS, W. O. Studies in the spread of effect. *J. comp. Psychol.*, 1943, **35**, 41–72.
9. JENKINS, W. O., AND SHEFFIELD, F. D. Rehearsal and guessing habits as sources of the "spread of effect." *J. exp. Psychol.*, 1946, **36**, 316–330.
10. MILLER, N. E., AND DOLLARD, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
11. MUELLER, C. G. Numerical transformations in the analysis of experimental data. *Psychol. Bull.*, 1949, **46**, 198–223.
12. MUENZINGER, K. F., DOVE, C. C., AND BERNSTONE, A. H. Serial learning: II. The bi-directional goal gradient in the endless maze. *J. genet. Psychol.*, 1937, **50**, 229–241.
13. NAGATY, M. O. The effect of reinforcement on closely following S–R connections: I. The effect of a backward conditioning procedure on the extinction of conditioned avoidance. *J. exp. Psychol.*, 1951, **42**, 239–246.
14. PAVLOV, I. P. *Conditioned reflexes* (G. V. Anrep, tr.). London: Oxford Univ. Press, 1927.
15. RUCH, F. L. Goal direction orientation, generalized training habit, and goal gradient as factors in maze learning in the rat. *J. comp. Psychol.*, 1934, **17**, 225–232.
16. SHEFFIELD, F. D. "Spread of effect" without reward or learning. *J. exp. Psychol.*, 1949, **39**, 575–579.
17. SMITH, M. H., JR. Spread of effect is the spurious result of non-random response tendencies. *J. exp. Psychol.*, 1949, **39**, 355–368.
18. SPENCE, K. W. The order of eliminating blinds in maze learning by the rat. *J. comp. Psychol.*, 1932, **14**, 9–27.
19. THOMPSON, M. E., AND DOVE, C. C. A further study on the bi-directional goal gradient in the endless maze. *J. exp. Psychol.*, 1943, **33**, 447–456.
20. THORNDIKE, E. L. A proof of the law of effect. *Science*, 1933, **77**, 173–175.
21. WALLACH, H., AND HENLE, M. An experimental analysis of the law of effect. *J. exp. Psychol.*, 1941, **28**, 340–349.
22. WALLACH, H., AND HENLE, M. A further study of the function of reward. *J. exp. Psychol.*, 1942, **30**, 147–160.
23. WASHBURN, M. F. *The animal mind.* (4th Ed.) New York: Macmillan, 1936.
24. ZIRKLE, G. A. Success and failure in serial learning: II. Isolation and the Thorndike effect. *J. exp. Psychol.*, 1946, **36**, 302–315.

EXPERIMENTS ON SENSORY-TONIC FIELD THEORY OF PERCEPTION: I. EFFECT OF EXTRANEOUS STIMULATION ON THE VISUAL PERCEPTION OF VERTICALITY [1]

BY SEYMOUR WAPNER, HEINZ WERNER, AND KENNETH A. CHANDLER

Clark University

The history of research in perception contains many studies which show the effect of organismic states, such as posture, motor activity, etc., upon visual space. Many psychologists have viewed these results as dealing with rather marginal aspects of perception because they emphasized visual processes as isolated events. It is only recently, owing to interest in the projective character of perception, that attention has turned to organismic states as important factors in perceptual organization.[2]

The latter view is central for the sensory-tonic field theory of perception (**2**). The theory postulates that the state of the organism is part and parcel of perceptual events, or specifically, that perception is the product of an interaction between object stimuli (sensory-tonic) and organismic state (somato- and viscerotonic). This postulate leads directly to specific hypotheses amenable to experimental investigation.

The particular hypothesis underlying the present study comes from the following considerations. According to the theory, with a change in organismic state, a change in perception is expected. One method of influencing the status of the organism is by "extraneous stimulation"; by this term we mean any stimulation to the organism which comes from a source other than the object perceived. It is thus hypothesized that extraneous stimulation influences the perception of an object.

The present study was designed to investigate the effect of two kinds of extraneous stimulation, electrical stimulation to the neck muscle and auditory stimulation, on the visual perception of verticality.

Apparatus And Procedure

Stimulus object.—In all situations tests were carried out in a dark room. The task was to adjust a luminescent rod which could be rotated in the fronto-parallel plane so that it appeared vertical. The luminescent rod, 39 in. long and 1 in. wide, was pivoted at its center, and its angular displacement was read off in fractions of a degree from a protractor.

The *S* sat erect throughout all trials. His head was held in position by an adjustable headrest attached to the chair on which he sat.

Extraneous stimulation.—Electrical stimulation by means of an inductorium was applied to either the left or right neck muscle (sternocleidomastoid). The stimulating electrode was embedded in a collar which fitted closely around *S*'s neck. The second electrode was attached to the left or the right hand, corresponding to the side of the body which received electrical stimulation. Intensity of stimulation (a setting of 9 on a Harvard Inductorium) was constant for all *S*s. The stimulation was continuous during each observation, which lasted for 10 sec.

Auditory stimulation was applied by means of two Western Electric headphones embedded in the headrest. An AC operated audio-oscillator provided an auditory stimulus of 800 cycles; sound intensity (11 db—1.73 v./500 ohms) was constant for all *S*s.

Test conditions.—In all conditions of testing, *S* reported the apparent position of the rod to *E* and informed him which way to move the rod to make it appear vertical. The *E* turned the rod

[1] This investigation was supported in part by a research grant from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

[2] See the recent symposium on the interrelationships between perception and personality (**1**).

slowly by increments of approximately 3° until *S* stated that the rod was vertical. The task was carried out under five conditions: (a) control—no stimulation; (b) auditory stimulation—left ear; (c) auditory stimulation—right ear; (d) electric stimulation—left; and (e) electric stimulation—right. Each *S* made four adjustments to the vertical under each of the experimental conditions.

The *S* had his eyes covered with goggles at all times except when adjusting the rod to the vertical. When given a signal, *S* removed his goggles, adjusted the rod to vertical, and replaced his goggles until the next trial. For those trials where extraneous stimulation was applied the procedure was as follows: At the onset of stimulation *S* raised his goggles. Stimulation continued until the final adjustment was made. When the extraneous stimulation ceased, *S* replaced his goggles. On half of the trials the initial setting of the rod was 30° left; on the other half, 30° right. The sequence of test conditions was: control, auditory-right, electrical-left, auditory-left, and electrical-right. A 1-min. rest period was given between each of the five conditions. A 20-sec. rest period was interposed between trials. Each trial lasted for approximately 10 sec.

Forty undergraduate students, 20 male and 20 female, were employed as *S*s.

Measures employed.—The position in which the rod looks vertical to *S* (apparent vertical) was measured in degrees of tilt with reference to objective vertical (plumb line). Thus all measures of the position of the luminescent rod were defined in terms of the perpendicular. Arbitrarily we designated an angular position clockwise (CW) of objective vertical (as viewed by *S*) as "+," and counterclockwise (CCW) as "−."

Results

General effects.—The data were analyzed by a two-way classification analysis of variance, with experimental condition as one variable and individual as the other. The analysis of variance and *t*-tests between mean scores for the test conditions are presented for men in Table I. The *F*-tests for experimental conditions and individuals are significant beyond the .01 level of confidence.

As can be seen from Table I, without stimulation (control), the apparent vertical is at a position of −1.4°, i.e., the luminescent rod is on the average perceived as vertical when it is rotated 1.4° CCW. With stimulation, however, the rod is a a position tilted to the left or right of the control depending on the side to which the stimulus is applied. The rod is to the left (CCW) of the control when stimulation is applied to the right side, and to the right (CW) when stimulation is applied to the left. The positions of the rod under all conditions of stimulation differ significantly from the position without stimulation. Furthermore, the positions under left stimulation differ significantly from the positions under right stimulation. The positions of the rod are identical for electrical and auditory extraneous stimulation when applied to the left side, and also identical for these two stimulations when applied to the right side.

In Table II the analysis of variance and *t*-tests are presented for women. The *F*-test for experimental conditions is significant at the .05 level and for individuals beyond the .01 level. Though the results are not as clear-

TABLE I

Significance of Difference of Means Among Test Conditions: Men

F-Tests

Source of Variance	Sum of Squares	*df*	*F*
Conditions (C)	377.36	4	7.26*
Subjects (S)	8378.16	19	33.95*
Residual (C × S)	987.04	76	
Total	9742.56	99	

t-Tests

Conditions	Mean	Mean Differences			
		A–L	A–R	E–L	E–R
No stimulus (C)	−1.4°	0.4*	0.7*	0.4*	0.7*
Auditory left (A–L)	−1.0°		1.1*	0.0	1.1*
Auditory right (A–R)	−2.1°			1.1*	0.0
Electric left (E–L)	−1.0°				1.1*
Electric right (E–R)	−2.1°				

* Significant at or better than the .01 level of confidence.

cut for women as they are for men, the general directional effects of stimulation still obtain. Stimulation on the right side results in rod position more to the left, and stimulation to the left side results in rod position more to the right. Auditory stimulation to the left side is significantly different from auditory and electrical stimulation to the right side.

Since the results for women were not as conclusive as those for men it appeared advisable to repeat the experiment. This was done twice for two separate groups of 20 women within a year after the original experiment was performed. The *F*-tests were significant for experimental conditions in each of the two repetitions of the experiment. As in the original experiment, there was a significant directional effect of extraneous stimulation as described above.

Individual differences.—It has already been noted from the analyses of variance that significant individual differences occurred for both men and women. The consistency of these differences was studied in two ways: within each condition by reliability coefficients, and between conditions by intercorrelations.

TABLE II

SIGNIFICANCE OF DIFFERENCE OF MEANS AMONG TEST CONDITIONS: WOMEN

F-Tests

Source of Variance	Sum of Squares	*df*	*F*
Conditions (C)	8.46	4	3.31*
Subjects (S)	97.73	19	8.03**
Residual (C × S)	48.26	76	
Total	154.45	99	

t-Tests

Conditions	Mean	Mean Differences			
		A–L	A–R	E–L	E–R
No stimulus (C)	−0.3	0.7*	0.0	0.2	0.0
Auditory left (A–L)	0.4		0.7*	0.5	0.7*
Auditory right (A–R)	−0.3			0.2	0.0
Electric left (E–L)	−0.1				0.2
Electric right (E–R)	−0.3				

* Significant at the .05 level of confidence.
** Significant at the .01 level of confidence.

TABLE III

RELIABILITY COEFFICIENTS

Condition	Men $N = 20$	Women $N = 20$
Control (no stimulus)	.83	.58
Auditory left	.94	.70
Auditory right	.84	.71
Electric left	.90	.74
Electric right	.80	.71

Table III presents uncorrected split-half reliability coefficients for men and women. All the correlations are significant beyond the .01 level of confidence. The coefficients are higher for men (.80 to .94) than for women (.58 to .74).

The intercorrelations among test conditions are shown in Table IV. All correlations are significant beyond the .01 level except two of the correlations for women which are significant at the .05 level. Here, as in the case of the reliability coefficients, corresponding correlations are lower for women than for men.

TABLE IV

INTERCORRELATIONS AMONG TEST CONDITIONS SEPARATELY FOR MEN AND WOMEN*

Conditions	A–L		A–R		E–L		E–R	
	M	*W*	*M*	*W*	*M*	*W*	*M*	*W*
No stimulus (C)	.94	.54	.82	.51	.92	.57	.92	.71
Auditory left (A–L)			.84	.62	.92	.83	.92	.64
Auditory right (A–R)					.82	.66	.80	.74
Electric left (E–L)							.84	.78
Electric right (E–R)								

* A correlation of .44 is significant at the .05 level, and .56 is significant at the .01 level of confidence.

Discussion

The most significant finding of this study is that visual perception of of verticality is affected by "extraneous stimulation"—in this instance, electrical stimulation of the neck muscle, and auditory stimulation. These results are in accordance with the general assumption of the sensory-tonic field theory, namely, that the state of the organism is an integral part of a perceptual event; specifically, if the stimulus object is visual, its perception will be affected by stimulation arising from non-visual sources. In particular, there is a specific effect of extraneous stimulation when asymmetrically applied: if an extraneous stimulus is applied to one side of the body a line appears vertical when rotated away from the side of the stimulation.

Another important finding pertains to the efficacy of the two types of extraneous stimulation. As noted in the results, the two kinds of extraneous stimulation used in this experiment function in a similar way. In general, both types of stimulation, applied to either of the two sides of the body, shift the apparent vertical similarly with respect to direction as well as amount. These findings are in accord with the notion of "functional equivalence," an essential construct within the sensory-tonic field theory of perception. Interaction of sensory and organismic events logically demands that factors which interact must have common dynamic properties; i.e., they must be equivalent with respect to their contribution to the total dynamic process. Therefore, one of the important implications of our findings is that they experimentally demonstrate such equivalence between diverse kinds of extraneous stimulation, the one ordinarily classified as auditory, the other as muscular.

Finally, one other result and its implication may be mentioned. The mean position of apparent vertical seen without extraneous stimulation deviates from the objective vertical (plumb line); and, most important, individuals are consistent in the way their apparent vertical deviates from objective vertical. This suggests that the individual organismic state, independent of extraneous stimulation, is an important variable which affects the position of the apparent vertical.[3] The significance of this variable for the interpretation of perceptual facts within our theoretical framework may become clearer from experimental evidence to be presented in forthcoming studies.

Summary

The present study was designed to investigate the effect of two kinds of stimulation from diverse sources—electrical stimulation to the neck muscle, and auditory stimulation—on the visual perception of verticality.

Forty *S*s, 20 men and 20 women, were used. The task for *S* consisted in the adjustment of a luminescent rod to a position of apparent verticality. It was found that asymmetrical stimulation affected the position of the apparent vertical significantly. Under stimulation to one side of the body the position of apparent vertical was shifted relatively to the side opposite the stimulation.

Electrical stimulation to the neck muscle and auditory stimulation affected perception of the vertical in an equivalent manner.

[3] The individual's characteristic position of apparent vertical indicates an organismic state of equilibrium which we refer to as "individual equilibrium."

Highly consistent individual differences were evident.

The results were interpreted within the framework of the sensory-tonic field theory of perception.

(Manuscript received for priority publication July 16, 1951)

References

1. Interrelationships between perception and personality: a symposium, Pt. I, *J. Personality*, 1949, **18**, 1–143; Pt. II, *J. Personality*, 1949, **18**, 145–266.
2. Werner, H., and Wapner, S. Sensory-tonic field theory of perception. *J. Personality*, 1949, **18**, 88–107.

EXPERIMENTS ON SENSORY-TONIC FIELD THEORY OF PERCEPTION: II. EFFECT OF SUPPORTED AND UNSUPPORTED TILT OF THE BODY ON THE VISUAL PERCEPTION OF VERTICALITY [1]

BY HEINZ WERNER, SEYMOUR WAPNER, AND KENNETH A. CHANDLER

Clark University

It is well known that with small sidewise tilts of the body a luminescent line in the dark will be perceived as vertical when tilted an appropriate number of degrees in a direction opposite to body tilt.[2] Most of the studies reporting this phenomenon have been made under conditions where the body is passively supported either by a chair or a board. The question arises whether there would be a differential effect if *S* had to maintain a specific body tilt unsupported by the side and back of a chair or a leaning board.

Within the framework of the sensory-tonic field theory of perception (**4**) the degree of muscular involvement is expected to be an important variable; therefore, differences in perception under supported and unsupported conditions might be predicted.

The theory suggests the following specific hypothesis: Under identical angles of body tilt, the position in which the rod looks vertical to *S* (apparent vertical) should be displaced to a greater degree from the perpendicular when the body is unsupported than when it is supported.

Since, in general, it can be assumed that with greater magnitude of tilt there is greater muscular involvement, a second hypothesis may be formulated: Whether supported or unsupported, with an increase of body tilt, there should be an increase in angular displacement of the apparent vertical.[3]

It is the purpose of this study to test these hypotheses.

METHOD

Stimulus object.—In all situations tests were carried out in a dark room. The task was to adjust a luminescent rod which could be rotated in the fronto-parallel plane so that it appeared vertical. The luminescent rod, 39 in. long and 1 in. wide, was pivoted at its center, and its angular displacement was read off in fractions of a degree from a protractor.

Experimental design.—Three variables were systematically studied: (a) kind of support—sitting fully supported and sitting relatively unsupported; (b) magnitude of tilt—15° and 30°; (c) direction of sidewise tilt—left and right. All combinations of these three variables were employed, making a total of eight experimental conditions: (1) 15° left—supported; (2) 30° left—supported; (3) 15° left—unsupported; (4) 30° left—unsupported; (5) 15° right—supported; (6) 30° right—supported; (7) 15° right—unsupported; and (8) 30° right—unsupported.

Forty *S*s, 20 male and 20 female, were tested under all eight experimental conditions.

For purposes of statistical analysis the data were separated according to sex and direction of tilt. Thus, analyses of variance evaluating the effect of kind of support and magnitude of tilt were carried out for four sets of data: (a) men—left body tilt; (b) men—right body tilt; (c) women—left body tilt; and (d) women—right body tilt.

Procedure.—Two major test conditions concerned the manner in which the *S* sat while making observations. Under conditions of being fully supported, *S* sat in a chair which was pivoted below the seat permitting left and right tilts.

[1] This investigation was supported in part by a research grant from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

[2] This is the so-called E-phenomenon. See Aubert (**1**), G. E. Müller (**2**), etc., and most recently Witkin and Asch (**5**).

[3] These hypotheses are restricted to the relatively small tilts in which the E-phenomenon has been generally observed.

The chair was provided with a headrest, sides and back; thus the body was held firmly in place. Under relatively unsupported conditions *S* sat on a stool, and was required to maintain his body in the desired inclined position by his own muscular efforts. To be certain that *S* assumed the designated angle of body tilt the following device was employed. A metal pointer was fixed to *S*'s back by means of waist and head straps; it extended above his head in line with his spinal column. The tip of the pointer was painted with luminescent paint. Behind the pointer was a luminescent grid marked off in degrees. The *E* directed *S* to tilt his body until the luminescent pointer coincided with the appropriate mark on the grid behind *S*. Throughout each trial *E* made certain that *S* maintained the desired angle of inclination.

In all conditions of testing, *S* reported the apparent position of the rod to *E* and informed him how to move the rod to make it vertical. The *E* turned the rod by small increments until *S* stated that the rod was vertical. The *S*'s eyes were covered with goggles at all times except when adjusting the rod to the vertical.

The sequence of the eight test conditions for all *S*s was: 15°—right, 15°—left, 30°—right, and 30°—left while supported, followed by the same sequence for the unsupported condition. Four trials were given for each experimental condition. Each trial started with the rod in a tilted position. On two trials, the initial setting of the rod was 30°—left; on the other two trials 30°—right.

Measures employed.—The position in which the rod looked vertical to *S* (apparent vertical) was measured in degrees of tilt with reference to objective vertical (plumb line). Thus all measures of the position of the luminescent rod were defined in terms of the perpendicular. Arbitrarily we designate an angular position clockwise (CW) of objective vertical (as viewed by *S*) as "+" and counterclockwise (CCW) as "−." The measure used to represent each *S*'s performance is the algebraic mean of the positions of the rod for the four trials of each test condition.

Results

The results are summarized in Table I. This table shows the mean positions of apparent vertical for each of the test conditions. For every test condition but one, the apparent vertical is tilted, with respect to objective vertical, in a direction opposite to body tilt. Furthermore, it can be seen that for identical angles of body tilt, the degree of deviation from objective vertical is greater for the unsupported than the supported condition. This relationship between supported and unsupported conditions exists whether the body is tilted to the left or to the right.

TABLE I

Mean Position of Apparent Vertical with Reference to Objective Vertical for Supported and Unsupported Body Tilt *

Condition	Body Tilt			
	Right		Left	
	15°	30°	15°	30°
Men ($N = 20$)				
Supported	−3.2°	−7.4°	+1.5°	+3.8°
Unsupported	−4.2°	−8.6°	+4.8°	+7.1°
Women ($N = 20$)				
Supported	−1.8°	−7.8°	−1.0°	+3.0°
Unsupported	−4.0°	−11.5°	+4.3°	+6.7°

* The plus sign indicates a position clockwise of objective vertical as viewed by *S*; the minus sign indicates a counterclockwise position.

It is also evident that within a given condition of support, the 30° body tilt results in a greater displacement of apparent vertical than the 15° body tilt.

Tests of significance of the differences between mean positions of the apparent vertical are presented in Table II. This table summarizes the analyses of variance done separately for sex and direction of body tilt. The *F*-test for supported versus unsupported conditions is significant in three out of the four analyses. In the one analysis where the *F*-test is not significant (Men—right body tilt), the relationship between the means is the same, i.e., the deviation of apparent vertical is greater under

TABLE II

ANALYSES OF VARIANCE

Source of Variance	*df*	Women						Men					
		Left Body Tilt			Right Body Tilt			Left Body Tilt			Right Body Tilt		
		Sum of Squares	*F*	*p*	Sum of Suqares	*F*	*p*	Sum of Squares	*F*	*p*	Sum of Squares	*F*	*p*
Kind of support (S)	1	404.55	44.46*	<.01	174.34	6.90*	<.05	222.44	9.41***	<.01	25.09	<1.00*	>.05
Magnitude of tilt (M)	1	201.93	22.19*	<.01	913.28	36.16*	<.01	109.04	5.26**	<.05	377.58	5.64**	<.05
Individuals (I)	19	948.53	5.49*	<.01	1341.96	2.80*	<.05	1012.15	2.25***	<.05	1902.12	1.50**	>.05
S × M	1	11.63	1.28*	>.05	10.01	<1.00*	>.05	0.03	<1.00*	>.05	0.31	<1.00*	>.05
S × I	19	248.51	1.44*	>.05	254.21	<1.00*	>.05	449.36	8.30*	<.01	225.52	<1.00*	>.05
M × I	19	299.71	1.73*	>.05	834.75	1.74*	>.05	394.14	7.28*	<.01	1272.19	4.26*	<.01
S × M × I	19	172.92			479.89			54.10			298.27		
Total	79	2287.78			4008.44			2241.23			4101.08		

* Tested against S × M × I.
** Tested against M × I.
*** Tested against S × I.

TABLE III

RELIABILITY COEFFICIENTS

Body Tilt	Men		Women	
	Supported	Unsupported	Supported	Unsupported
15° left	.97	.77	.85	.92
30° left	.90	.84	.95	.87
15° right	.94	.88	.91	.96
30° right	.97	.93	.91	.82

unsupported than supported conitions.

As to the degree of body tilt, the *F*-tests are significant in every instance. This indicates that the deviation of the apparent vertical from objective vertical is significantly greater at a 30° inclination of the body than at a 15° inclination.

In Table III the split-half uncorrected reliability coefficients are presented for men and women. All of the correlations are significant and high. Thus, within each test condition the individuals respond in a strikingly stable manner.

In summary, the empirical data obtained confirmed the hypotheses formulated in this study: (a) Under identical angles of body tilt the position of the rod indicating apparent verticality is displaced to a greater degree from the perpendicular when the body is unsupported than when it is supported. (b) Irrespective of kind of support, angular displacement of the apparent vertical increases with an increase of body tilt.

DISCUSSION

If one scans the literature on the perceptual displacement of the vertical under body tilts (e.g., E-phenomenon), one finds few or no interpretative attempts that go beyond the rather vague statement that postural factors and their concomitants underlie this effect. Such interpretations, even when they are couched in terms of specific changes due to posture (such as kinaesthetic, cutaneous, oculomuscular, retinal-positional, etc.), are confronted with two theoretical difficulties.

One has to do with the concept of interaction. Most theorists simply state that interaction of two processes (e.g., kinaesthetic and visual) occurs, and fail to see the logical necessity of providing a framework within which these otherwise alien interacting elements can be conceived as dynamically equivalent. Whatever the specific explanations of the E-phenomenon may be, we believe that the sensory-tonic field theory of perception provides, at least, a logical basis by which the mechanism of interaction can be understood.

The second difficulty pertains to the notion of posture as related to our results. It was found that the degree of perceptual displacement depends upon kind of support (amount of muscular involvement) rather than simply on the angle of body tilt per se. In the light of these results, any theory which would try to understand the E-phenomenon by a static concept of body position must fail. By the same token, explanations in terms of derivatives of static postural states, such as specific retinal position, labyrinthian reflexes, etc., are equally unsatisfactory.

In contrast, the sensory-tonic field theory of perception rejects the conception of posture as a static position of the body; it identifies posture as a dynamic state of the organism that involves the continuous patterning of somato- and viscerotonic processes. The crucial postural factor is not position per se but degree of muscular involvement. It is this factor which partakes of the sensory-tonic field,

and thus contributes to changes in perception.

This dynamic conception of posture makes it possible to integrate the effects of extraneous stimulation, as presented in a previous study (**3**), with the results on posture, as discussed here. It was demonstrated there that auditory stimulation and electrical stimulation of the neck muscle, asymmetrically applied, displace the apparent vertical in a direction opposite to the side of stimulation. These effects are analogous to the effects of body tilt though in this case the body is erect. We are thus justified to speak of an equivalence between postural and nonpostural (extraneous stimulation) factors with reference to effects on perception. This leads us again back to the interpretation made above that the efficacy of posture resides in the underlying somatotonic processes rather than in body position per se.

Summary

The present study was designed to investigate the effect of postural changes upon the visual perception of verticality. Apparent verticality was studied while *S*'s body was tilted at specified angles. Observations were made under two main conditions: (a) supported—*S* was supported by the sides and back of a chair; and (b) unsupported—*S* had to maintain the angle of tilt by his own muscular efforts while sitting on a stool.

Forty *S*s, 20 men and 20 women, were tested. The task for *S* consisted in the adjustment of a luminescent rod to a position of apparent verticality. The main results are: (a) the apparent vertical was shifted to the side opposite the direction of body tilt; (b) with increase of body tilt from 15° to 30° there was a significantly greater shift of the apparent vertical; and (c) under identical angles of body tilt, the apparent vertical was displaced to a greater degree when the body was unsupported than when it was supported.

These results are in accordance with specific hypotheses derived from the sensory-tonic field theory of perception.

(Manuscript received for priority publication July 30, 1951)

References

1. Aubert, H. Eine scheinbare bedeutende Drehung von Objecten bei Neigung des Kopfes. *Virchow's Arch.*, 1861, **20**, 381–393.
2. Müller, G. E. Über das Aubert'sche Phänomen. *Z. Psychol.*, 1916, II, **49**, 109–244.
3. Wapner, S., Werner, H., and Chandler, K. A. Experiments on sensory-tonic field theory of perception: I. Effect of extraneous stimulation on visual perception of verticality. *J. exp. Psychol.*, 1951, **42**, 341–345.
4. Werner, H., and Wapner, S. Sensory-tonic field theory of perception. *J. Personality*, 1949, **18**, 88–107.
5. Witkin, H. A., and Asch, S. E. Studies in space orientation. III. Perception of the upright in the absence of a visual field. *J. exp. Psychol.*, 1948, **38**, 603–614.

EXPERIMENTS ON SENSORY-TONIC FIELD THEORY OF PERCEPTION: III. EFFECT OF BODY ROTATION ON THE VISUAL PERCEPTION OF VERTICALITY [1]

BY SEYMOUR WAPNER, HEINZ WERNER, AND RICARDO B. MORANT

Clark University

It has been demonstrated in previous papers (**9, 13**) that a change in the state of the organism significantly affects the visual perception of verticality. In the first of the studies the method of "extraneous stimulation" was used to influence the organismic state. The term "extraneous stimulation" was there defined as stimulation to the organism coming from a source other than the object perceived. Two kinds of extraneous stimulation were employed, *viz.*, electrical stimulation to the neck muscle, and auditory stimulation.

The present experiment was designed to study the effect of another kind of extraneous stimulation on visual perception, i.e., labyrinthian stimulation induced by rotation of the body around its vertical axis.

Apparatus And Procedure

Stimulus object.—The visual stimulus employed consisted of a luminescent rod 39 in. long and 1 in. wide. The rod was pivoted at its center and could be rotated in the fronto-parallel plane. The *S* had the task of adjusting this rod to the vertical while rotating in a chair. The apparatus was arranged so that the visual stimulus moved with *S*, i.e., was always in front of him; the movement of the rod was remotely controlled by *E*.

Remote control.—The rod was driven by a commercially built antenna rotor. Since remote control of rod position by *E* had to be achieved during the spinning of *S*'s chair, commutator rings encircling the supporting shaft of the chair provided the electrical connection between the rotor and *E*'s desk. The rotor was connected by a chain drive to a selsyn actuator which in turn acted in phase with a selsyn repeater at *E*'s desk. This selsyn system provided an indication of the position of the rod in half degrees. Thus, at the direction of *S*, *E* was able to manipulate the luminescent rod remotely while the chair was in motion, and to read rod position directly by means of the selsyn system.

The rod, together with the rotor and selsyn actuator, was attached to the rotating chair by an iron pipe. The pipe extended from the lower part of the chair horizontally and was bent sharply upwards at a distance of 5 ft. in front of *S*, so that the center of the rod was at eye level (see Fig. 1).

Rotating chair.—The rotating chair was supported by a metal shaft mounted on ball bearings and carefully balanced to assure a smooth rotary motion. The rotation was controlled by a constant speed motor through a Graham Variable Speed drive. One speed, 25 rpm, was used for both clockwise (CW) and counterclockwise (CCW) motion. The period of acceleration, from zero to a constant speed of 25 rpm, lasted 15 sec. The period of deceleration, from constant speed of 25 rpm to zero, lasted 9 sec.

The head was fixed in position by means of a head rest attached to the chair. The head was tipped forward so that a line from the corner of the eye to the ear canal was horizontal. Thus, the horizontal canals were approximately in the horizontal position during testing.

Test conditions.—All testing was carried out in the dark. Between trials *S*'s eyes were closed and covered by opaque rubber rimmed goggles. During a trial when *S*'s eyes were open, only the luminescent rod was visible to him.

Five test conditions were used: (a) control—absence of rotation (C); (b) acceleration in a clockwise direction (A–CW); (c) deceleration from constant clockwise rotation (D–CW); (d) acceleration in a counterclockwise direction (A–CCW); (e) deceleration from constant counterclockwise rotation (D–CCW).

For half of the *S*s, the sequence of conditions was as follows: a-b-c-rest-a-d-e-rest-a-d-e-rest-a-b-c. For the other half of the *S*s the sequence was: a-d-e-rest-a-b-c-rest-a-b-c-rest-a-d-e. The rest period, which always preceded the control

[1] This investigation was supported by a research grant from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

FIG. 1. Rotating chair and control desk

condition, lasted for 5 min.[2] Two consecutive measurements were taken for each presentation of a test condition. Thus a total of eight measurements were taken for control, and four measurements for each of the four rotation conditions.

Twenty-eight *S*s, 14 men and 14 women, were tested.

Measures employed.—The position in which the rod looked vertical to *S* (apparent vertical) was measured, with reference to objective vertical, in degrees of tilt. Therefore, all measures of the position of the luminescent rod were defined in terms of the perpendicular. Arbitrarily, we designate an angular position clockwise (CW) of objective vertical as "+" (tilted right) and counterclockwise (CCW) as "−" (tilted left). The measure used to represent each *S*'s perception is the algebraic mean of the adjustments for each condition.

RESULTS

The results are summarized diagrammatically in Fig. 2. The figure shows the mean position of apparent vertical for control and for each of the test conditions, separately for men and women. Under CW acceleration, the apparent vertical is tilted to the left of control; under CCW acceleration the apparent vertical is tilted to the right of control. Under deceleration from CW rotation, the apparent vertical is tilted to the right of control; under deceleration from CCW rotation the apparent vertical is tilted to the left of control.

Thus, the apparent vertical is shifted opposite the direction of acceleration, i.e., when acceleration is to the right, the tilt is to the left and vice versa.[3] Moreover, deceleration from rotation in one direction functions like acceleration in the opposite direction, i.e., in deceleration from rotation to the right, apparent vertical is tilted to the same side as in acceleration to the left, and vice versa.

[2] There was no significant differential effect due to sequence.

[3] This statement is dynamically but not geometrically correct. That is, CW and CCW with reference to rotation deal with motion around the vertical axis of the body; whereas, with reference to the rod, CW and CCW deal with motion in the fronto-parallel plane.

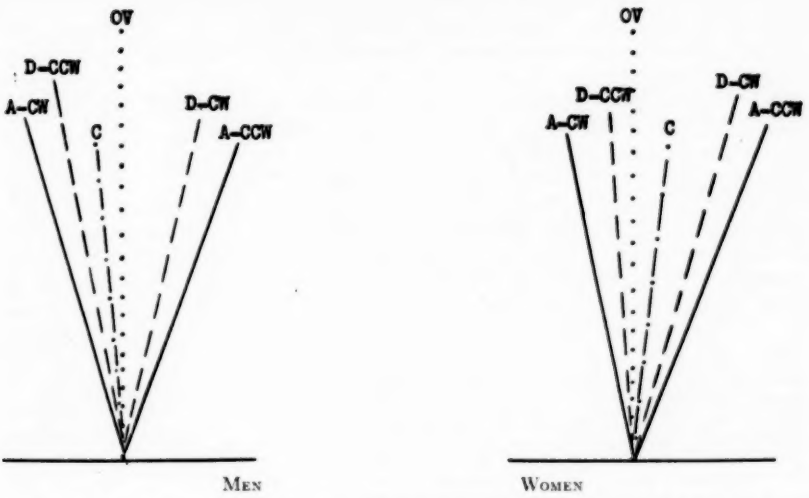

FIG. 2. Mean position of apparent vertical with reference to objective vertical (OV) under labyrinthian stimulation (effect magnified ten times): A–CW—clockwise acceleration; D–CCW—deceleration from counterclockwise rotation; C—control; D–CW—deceleration from clockwise rotation; A–CCW—counterclockwise acceleration.

Tests of significance of the differences between mean positions of apparent vertical shown in Fig. 2 were made by a two-way classification analysis of variance separately for men and women. The results of these analyses are summarized in Tables I and II.

TABLE I

SIGNIFICANCE OF DIFFERENCE OF MEANS AMONG TEST CONDITIONS: MEN

F-Tests

Source of Variance	*df*	Sum of Squares	*F*
Conditions (C)	4	2178.75	33.58**
Subjects (S)	13	554.49	2.63**
Residual (C × S)	52	843.21	
Total	69	3576.45	

t-Tests

Conditions	Mean	Mean Differences			
		A–CW	D–CW	A–CCW	D–CCW
Control (C)	−0.1	1.5**	1.6**	2.0**	1.1*
Acceleration CW (A–CW)	−1.6		3.1**	3.5**	0.4
Deceleration from CW rotation (D–CW)	+1.5			0.4	2.7**
Acceleration CCW (A–CCW)	+1.9				3.1**
Deceleration from CCW rotation (D–CCW)	−1.2				

* Significant at or better than the .05 level of confidence.
** Significant at or better than the .01 level of confidence.

TABLE II

SIGNIFICANCE OF DIFFERENCE OF MEANS AMONG TEST CONDITIONS: WOMEN

F-Tests

Source of Variance	*df*	Sum of Squares	*F*
Conditions (C)	4	1498.50	17.38**
Subjects (S)	13	1499.80	5.35**
Residual (C × S)	52	1121.22	
Total	69	4119.52	

t-Tests

Conditions	Mean	Mean Differences			
		A–CW	D–CW	A–CCW	D–CCW
Control (C)	+0.3	1.4**	1.3**	1.7**	0.6
Acceleration CW (A–CW)	−1.1		2.7**	3.1**	0.8
Deceleration from CW rotation (D–CW)	+1.6			0.4	1.9**
Acceleration CCW (A–CCW)	+2.0				2.3**
Deceleration from CCW rotation (D–CCW)	−0.3				

** Significant at or better than the .01 level of confidence.

The *F*-tests for test conditions and Ss are significant beyond the .01 level for both men and women. The mean position of apparent vertical without stimulation differs significantly from the apparent vertical under each of the experimental test conditions except one (Women, D-CCW). In every instance acceleration in one direction differs significantly from acceleration in the other direction; this holds also for deceleration. Acceleration in one direction differs significantly from deceleration from rotation in the same direction. In contrast, acceleration in one direction does not differ significantly from deceleration from rotation in the other direction.

Table III presents uncorrected reliability coefficients for each test condition separately for men and women. Six of the correlations are significant at the .01 level, three are significant at the .05 level, and one (D-CW, Men) only approaches significance.

TABLE III

UNCORRECTED RELIABILITY COEFFICIENTS

Conditions	Men	Women
C	.83**	.95**
A–CW	.59*	.59*
D–CW	.52	.70**
A–CCW	.61*	.86**
D–CCW	.85**	.73**

* Significant at the .05 level of confidence.
** Significant at or better than the .01 level of confidence.

DISCUSSION

The most important finding of this study is that visual perception of verticality is significantly affected by accelerative and decelerative rotation around *S*'s vertical axis. Specifically, the apparent vertical under CW acceleration is tilted to the left and under CCW acceleration is tilted to the right of control. Deceleration from rotation in one direction functions like acceleration in the opposite direction.

Studies which have dealt with the effect of labyrinthian stimulation on perception of the vertical date back

to Purkinje (**8**), Mach (**5**), and Breuer and Kreidl (**1**).[4] These older studies were not concerned with rotation around *S*'s own vertical axis (concentric), but with rotation around a circular path (eccentric). Only one older study, by Brünings (**4**), dealt with concentric rotation; this author, however, did not find a measurable effect on verticality. From Brünings' work and his own studies on eccentric rotation, Fischer (**4**, p. 267) erroneously concluded that eccentricity of rotation was a necessary condition for angular shifts in apparent verticality. Recent studies dealing with angular shifts of apparent vertical under labyrinthian stimulation, such as those by Noble (**7**) and Witkin (**14**), are confined to observations made under eccentric rather than concentric motion.

In order to explain such labyrinthian effects as studied here, two types of internally consistent theories are possible. One type concerns sensory theories interpreting the effects in terms of retinal positions. Another more inclusive type of theory takes into account sensory-tonic interactions. On the other hand, interpretations which attempt to explain the effects in terms of a correspondence between perceptual results and non-visual processes, kinaesthetic, labyrinthian changes, etc., are faced with the "interaction paradox." By "interaction paradox" we understand the contradiction inherent in the attempt to fabricate a unity out of alien elements, such as visual on one side and non-visual on the other.

It is unlikely that a purely sensory theory is capable of accounting for the facts. Breuer and Kreidl (**1**), among others, suggested changes of retinal position through torsion as the explanation for shifts of the apparent vertical under eccentric rotation. However, as recent studies have shown (**7**, **14**), under conditions of high speed eccentric rotation, perceptual shifts occur of such great magnitude that explanations in terms of torsion are highly improbable. There is another more basic reason why a retinal theory might be rejected. This deals with the concept of maximum generality of any theory. A retinal theory may cover the facts under very special sensory conditions; but it cannot account for analogous shifts of the apparent visual vertical observed under extraneous stimulation which most likely excludes eyeball torsion.[5] It is even more difficult for this kind of theory to account for analogous effects which have been observed in non-visual situations: here the apparent vertical was presented kinaesthetically with closed eyes (**11**), and effects occurred which were similar to those found with the visual vertical. Thus, to interpret the analogous results obtained in these diverse situations a theory is required which goes beyond the explanation of specific effects.

The sensory-tonic field theory of perception appears to fulfill the generality requirement. An explanation in terms of such a theory is greatly strengthened by previous extensive experimental evidence that the primary effect of labyrinthian impulses consists in their influence on the tonicity of musculature. Ewald (**3**) was probably the first to put forward the view that the constant flow of

[4] For a more complete summary of the effect of labyrinthian stimulation on perception see Dusser de Barenne (**2**), Fischer (**4**), and Wendt (**10**).

[5] We refer here to a previous paper (**9**) demonstrating the effect of extraneous stimulation (in that case auditory stimulation and electrical stimulation to neck muscles) on the position of apparent vertical.

impulses from the labyrinthian organs serves to maintain the muscles in normal condition of tonus. The well-known experiments by Magnus and de Kleyn (**6**) have given strong support for, and extension of, this view. Thus, the sensory-tonic field theory of perception identifies labyrinthian stimulation—like extraneous stimulation (auditory, etc.) and direct postural changes—as affecting and thus contributing to the dynamic status of the organism. It is the status of the organism, conceived of as a continuous patterning of somato- and viscerotonic processes, which, as a principal factor in the sensory-tonic field, makes for perceptual shifts of an otherwise enduring stimulus object.

It may be further noted that a common feature of the effects of auditory stimulation, electrical stimulation on neck musculature, postural tilts, and labyrinthian stimulation on verticality pertains to the direction of perceptual shifts. The previous studies on the effect of extraneous stimulation, and postural tilts have demonstrated opposite (CW-CCW) perceptual shifts corresponding to opposite (left-right) applications of stimuli or tilts of the body. It has been argued there that the factors responsible for these shifts in both situations are the induced asymmetrical tonic changes. Since opposite labyrinthian stimulations are involved in the present experimental situation (CW versus CCW acceleration, etc.) yielding opposite tonic effects, corresponding perceptual shifts of apparent vertical are expected.

Summary

The present study was designed to investigate the effect of labyrinthian stimulation induced by rotation of the body around its vertical axis (concentric) on the visual perception of verticality.

Twenty-eight *S*s, 14 men and 14 women, were used. The task for *S* consisted in the adjustment of a luminescent rod to apparent verticality under five test conditions: control—absence of rotation; clockwise and counterclockwise acceleration; deceleration from constant clockwise and counterclockwise rotation.

It was found that visual perception of verticality is significantly affected by accelerative and decelerative rotation around *S*'s vertical axis. The apparent vertical under clockwise acceleration is tilted in a counterclockwise direction, and under counterclockwise acceleration is tilted in a clockwise direction. Deceleration from rotation in one direction affects the apparent vertical like acceleration in the opposite direction.

These results are interpreted within the framework of the sensory-tonic field theory of perception.

(Manuscript received for priority publication August 10, 1951)

References

1. Breuer, J., and Kreidl, A. Ueber die scheinbare Drehung des Gesichtsfeldes während der Einwirkung einer Centrifugalkraft. *Pflügers Arch. ges. Physiol.*, 1898, **70**, 494–507.
2. Dusser de Barenne, J. G. The labyrinthine and postural mechanisms. In C. Murchison (Ed.), *Handbook of general experimental psychology*. Worcester: Clark Univ. Press, 1934.
3. Ewald, R. *Physiologische Untersuchungen über das Endorgan des Nervus Octavus*. Wiesbaden: Bergmann, 1892.
4. Fischer, M. H. Die Regulationsfunktionen des menschlichen Labyrinthes und die Zusammenhänge mit verwandten Funktionen. *Ergebn. Physiol.*, 1928, **27**, 209–379.
5. Mach, E. *Grundlinien der Lehre von den Bewegungsempfindungen*. Leipzig: Engelmann, 1875.

6. MAGNUS, R., AND DE KLEYN, A. Die Abhängigkeit des Tonus der Extremitätenmuskeln von der Kopfstellung. *Pflügers Arch. ges. Physiol.*, 1912, **145**, 455–548.
7. NOBLE, C. The perception of the vertical: III. The visual vertical as a function of centrifugal and gravitational forces. *J. exp. Psychol.*, 1949, **39**, 839–850.
8. PURKINJE, J. Beiträge zur näheren Kenntniss des Schwindels aus heautognostischen Daten. *Med. Jb. d. österreich. Staates*, 1820, **6**, 79–125.
9. WAPNER, S., WERNER, H., AND CHANDLER, K. A. Experiments on sensory-tonic field theory of perception: I. Effect of extraneous stimulation on visual perception of verticality. *J. exp. Psychol.*, 1951, **42**, 341–345.
10. WENDT, G. R. Vestibular functions. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
11. WERNER, H., AND WAPNER, S. Experiments on sensory-tonic field theory of perception: V. Effect of body and head tilt on kinaesthetic perception of verticality. (To be published.)
12. WERNER, H., AND WAPNER, S. Sensory-tonic field theory of perception. *J. Personality*, 1949, **18**, 88–107.
13. WERNER, H., WAPNER, S., AND CHANDLER, K. A. Experiments on sensory-tonic field theory of perception: II. Effect of supported and unsupported tilt of the body on the visual perception of verticality. *J. exp. Psychol.*, 1951, **42**, 346–350.
14. WITKIN, H. A. Perception of the upright when the direction of the force acting on the body is changed. *J. exp. Psychol.*, 1950, **40**, 93–106.

THE EFFECT OF PRESENTING VARIOUS NUMBERS OF DISCRETE STEPS ON SCALE READING ACCURACY [1]

BY HAROLD W. HAKE AND W. R. GARNER [2]

The Johns Hopkins University

In a recent paper (**3**) we demonstrated that the methodology of information theory can be applied usefully to the method of absolute judgments to answer two important questions. It can be used to determine how much information an *O* obtained about which particular stimulus was presented at each point in time; and the method gives a direct measure of the minimum number of stimuli from a set of stimuli which can transmit the maximum information about some other set of events.

The act of interpolating pointer positions between scale marks on an instrument dial is a judgment like that required in the method of absolute judgments. The *O* is not explicitly required to make a relative judgment. Rather, he is expected to identify each position with a single number. In addition, the successive pointer positions are code stimuli representing events on another continuum such as temperature or speed. For these reasons some of the experimental questions which researchers have attempted to answer in studying the perception of dials and scales are those which the methodology of information theory can answer.

In this experiment the method was applied to determine the minimum number of different pointer positions which can be presented in a standard interpolation interval to transmit the maximum amount of information, not about which positions of the pointer are occurring, but about the event continuum being represented.

We have shown (**3**) that the measure of amount of information indicates directly the minimum number of stimuli which can be used to transmit the maximum amount of information about a given set of possible events where each stimulus perfectly represents a unique set of these events. In general, the larger the errors of observation, the smaller this minimum number will be. Since the errors of observation in scale interpolation have been shown to be rather large and complex (**2, 4**), we would predict that the number of pointer positions which can be used in an interpolation interval on a scale to transmit maximum information about another variable is very much smaller than the number of different positions which the pointer on the average instrument can reliably assume.

A restriction in the number of different possible pointer positions, of course, can be thought of as resulting in instrument error because each pointer position then represents a range of values of the represented variable. In the course of indicating values in the interval, the restricted pointer could indicate the true value of the represented variable only when this value happened to correspond to one of the permissible positions. The information theory context predicts, however, that if each range of values

[1] This work was supported under Contract N5-ori-166, Task Order I, between the Special Devices Center, Office of Naval Research, and The Johns Hopkins University. This is Report No. 166-I-128, Project Designation No. NR-784-001, under that contract.

[2] Mr. Allan Thompson assisted in obtaining and analyzing the data of this experiment.

of the represented variable is perfectly represented by just one pointer position in the interpolation interval, then the amount of information transmitted will tend to be invariant provided that the number of pointer positions used is equal to or larger than some minimum number. This study provides experimental evidence demonstrating the validity of this statement.

Method

Apparatus.—The *S* was seated in a booth 39 in. long, 56 in. high, and 32 in. wide. He rested his head against a forehead rest and looked directly ahead at a ground glass plate 20 in. away and normal to the line of sight. During a trial, a rectangular lighted area .20 in. by .07 in. appeared in the center of the glass. An interpolation interval within the lighted area .16 in. long in the horizontal dimension was marked off by two dark vertical scale markers each .005 in. wide extending up from the bottom of the area. A vertical pointer, also .005 in. wide, extended down from the top of the area and indicated values within the interval. No numbers appeared anywhere on the glass plate. The illumination of the background and interval was such that the scale markers and pointer were all easily seen by *S*. The interpolation interval which *S* saw was the image of a larger interval located 10 ft. away from the booth and focused on the ground glass by a lens system. The position of the large pointer, the image of which *S* saw, could be set very precisely by *E*. *E* also controlled the floodlamps which illuminated the markers and pointer from behind, such that their silhouette was thrown on the glass in front of *S* during the duration of a trial.

Procedures.—At the beginning of an experimental session *S* looked directly ahead at the interpolation interval in which the pointer indicated the value "40." He was told that the interval represented a small unmarked interval which had been removed from a larger scale consisting of many such intervals. It was explained that the left marker represented the value "0," that the right marker represented the value "100," and that the pointer indicated the value "40" in this interval.

Four different series of pointer positions were then presented to *S*. In one of these series (five positions) only six different positions of the pointer were permitted. These were at 0, 20, 40, 60, 80, and 100. The positions at "0" and "100" were considered identical in all series for purposes of analysis because, normally on a scale, the value "0" in our interpolation interval would represent the highest value of the next lowest interval. Since the "0" and "100" positions were presented just one-half as often as the others, their combination provided five positions all presented equally often. The other three series, with the first and last positions combined, were composed of pointer positions at every tenth integer beginning with 10 (10 positions), every fifth integer beginning with 5 (20 positions), and every even integer beginning with 2 (50 positions).

The four series were presented under two conditions of instructions. If a series were to be read under the Limited Response condition, he was told the values of each permissible pointer position and that he must report each position only in terms of these values. If a series were to be read under the Unlimited Response condition, he was told to assume that the pointer could take any position within the interval and to report each position as accurately as he could, using the integers from zero to 100.

The order in which the four series were given to each *S* was determined by a modified Graeco-Latin square design in which each *S* read each pointer series in a predetermined random order, but he did not read every series under both conditions of instructions.

Each series of pointer positions consisted of 200 trials in which the positions occurred in chance order. At the beginning of each trial *S* was warned that a pointer position would be presented. The scale marks and pointer then appeared for 10 sec. during which time *S* reported the numerical value of the position. The intertrial interval was 10 sec.

Subjects.—The *S*s were 16 male undergraduates who had little or no experience in scale and dial reading. They were assigned to the experimental schedules by random assortment.

Results

Primarily we are interested in the discrepancies between the final readings made by *S* and the values on the original continuous variable, when the values on the original variable have been represented as discrete values. Many different measures are available to describe the discrepancies, but we have chosen the mean square error (variance) because the additivity characteristic of variances makes our calculations much simpler. We

have used the term "mean square error" rather than the term "variance" because the former measure includes constant errors as well as variable errors around a measure of central tendency.

The error analysis.—The error analyses are shown in Table I. The total mean square error represents the discrepancies between the final readings and the values of the continuous variable being represented, and is defined as:

$$\sigma_T^2 = \frac{1}{wi} \sum_{X=1}^{w} \sum_{i=1}^{i} (Y_i - X)^2 \quad (1)$$

where X represents the various possible values on the original continuum and Y_i represents the several readings made when a particular value of X occurred on the original continuum. Since it is impossible to know the value of X for any given response (since groups of X-values are represented as a single value on the scale presented to S), this formula cannot be used for a direct calculation. Fortunately, however, the formula can be broken into parts which can be separately computed. First, this total mean square error can be separated into two parts, one of which represents the discrepancies between the readings of S and the pointer positions; the other of which represents the discrepancies between the pointer positions and the values on the original continuum. Thus

$$\sigma_T^2 = \frac{1}{ki} \sum_{Z=1}^{k} \sum_{i=1}^{i} (Y_i - Z)^2 + \frac{1}{ki} \sum_{Z=1}^{k} \sum_{i=1}^{i} (X_i - Z)^2 \quad (2)$$

where the first part of the right-hand side of the equation is the Total Observer Error and the second part is the error due to the rounding off of values on the original continuum to discrete values (Instrument Error). In this equation, Z represents the various discrete values presented to S, and Y_i the readings made when a particular Z is presented. X_i represents the various values of X which are represented by a particular discrete value Z. The observer error can be directly calculated. The instrument error, however, must be evaluated as an integral, because the represented X variable is continuous.[3]

The total observer error can, in turn, be separated into two parts (**1**), as follows:

$$\sigma_O^2 = \frac{1}{ki} \sum_{\bar{Y}=1}^{k} \sum_{i=1}^{i} (Y_i - \bar{Y})^2 + \frac{1}{k} \sum_{Z=1}^{k} (\bar{Y} - Z)^2 \quad (3)$$

where $\bar{Y}$ is the mean reading made for a particular pointer position Z.

Table I shows not only the total mean square error, but the various components of the error, for each number of pointer positions and each condition of response limitation. The units in the error analysis are the values represented by the interpolation interval presented to the Ss. There are several facts which seem clear from this error analysis.

1. The instrument error increased as the square of the coarseness of the grouping, and when only five pointer positions were presented (the most coarse grouping) the instrument error was by far the largest part of the total error. With only five pointer positions, the instrument error was so large that no possible change in per-

[3] This integral is evaluated as $(X_2 - X_1)^2/12$, where X_2 and X_1 are the end points of the represented range of X, the pointer position corresponds to the mid-point of the range, and all values of the original variable are equally likely. This is essentially Sheppard's correction for grouping.

TABLE I

THE ERROR ANALYSIS *

Errors	Number of Positions			
	5	10	20	50
Limited Response				
Total Observer	.31	3.35	11.19	15.98
Variable	*.30*	*2.72*	*5.04*	*6.14*
Constant	*.01*	*.63*	*6.15*	*9.84*
Instrument	33.33	8.33	2.08	.33
Total	33.64	11.68	13.27	16.31
Unlimited Response				
Total Observer	15.77	13.16	20.14	11.72
Variable	*5.05*	*6.18*	*9.43*	*5.19*
Constant	*10.72*	*6.98*	*10.71*	*6.53*
Instrument	33.33	8.33	2.08	.33
Total	49.10	21.49	22.22	12.05

* The measure of error is the mean square error.

ceptual precision could compensate for it.

2. When Limited Responses were used, the total error was roughly constant when more than five pointer positions were used. The mean square error for observers decreased enough to compensate for the increased instrument error.

3. When Unlimited Responses were allowed, the observer error was relatively independent of the number of pointer positions. Since the instrument error increases with coarser grouping, the total error must also increase. Thus, the relative disadvantage of allowing unlimited responses increased as fewer pointer positions were used.

4. The total observer error is composed roughly of equal amounts of constant error and variable error, except when only five or ten pointer positions were used with Limited Responses. In these cases, the constant error is very small compared to the variable error. It seems likely that forcing *S* to use only responses which correspond to the permissible pointer positions prevents systematic deviations of his readings from the presented position.

Thus, we can conclude that there is no necessary loss in scale reading accuracy even though all possible scale values are represented by a few discrete values. In order for there to be no loss in scale reading accuracy, it is necessary for the reader of a scale to know how many possible pointer positions there are, and to be restricted to the use of responses which correspond to the presented discrete values.

The information analysis.—In this section the question of how much information the *S*s received about the coded information presented in the four numbers of pointer positions and under the two conditions of response limitation is answered. The error analysis presented above has indicated the relative efficiency of the experimental conditions, but the information analysis makes more explicit the important factors involved in the efficiency of transmission. There are four important components in this analysis. These are the Stimulus Information, the Response Information, the Response Equivocation, and

TABLE II

THE INFORMATION ANALYSIS *

Analysis	Number of Positions			
	5	10	20	50
Limited Response				
Stimulus Information	2.32	3.32	4.32	5.64
Response Information	2.32	3.32	4.29	5.53
Response Equivocation	.01	.18	1.13	2.34
Information Transmitted	2.31	3.14	3.16	3.19
Unlimited Response				
Stimulus Information	2.32	3.32	4.32	5.64
Response Information	4.79	5.49	5.99	5.98
Response Equivocation	2.50	2.46	2.88	2.57
Information Transmitted	2.29	3.03	3.11	3.41

* The amounts of information are in bits per appearance of the pointer.

the Information Transmitted. The method used in computing these components is Method (1) of the previous paper (**3**), and the results are presented in Table II.

The Stimulus Information for each number of pointer positions and each response limitation is presented in lines one and five of the table. This is the amount of information in bits about the represented variable which is available at each presentation of a pointer position. It can be thought of as the amount of uncertainty involved in the prediction of the next pointer position or the amount of information needed to predict it. The amount of Stimulus Information always sets an upper bound for the amount of information which the *S*s can obtain about the represented variable by watching the pointer. It is computed by

$$I_s = -\sum_{k=1}^{S} p(k) \log_2 p(k) \qquad (4)$$

where $p(k)$ is the probability of occurrence of position, k.

The amounts of Response Information are given in lines two and six of the table. They are computed by

$$I_r = -\sum_{j=1}^{R} p(j) \log_2 p(j) \qquad (5)$$

where $p(j)$ is the probability of S reading j irrespective of the particular pointer position presented. With limited responses, I_r tends to be equal or slightly less than I_s indicating that the *S*s tend to use all of the permissable responses about as often as the corresponding pointer positions are presented. With unlimited responses, I_r exceeds I_s by a considerable amount because the *S*s tend to respond with a wide variety of readings to each pointer position.

The Response Equivocation in lines three and seven is computed by

$$E_r = -\frac{1}{S}\sum_{k=1}^{S}\sum_{j=1}^{R} p(k)p_k(j) \log_2 p_k(j) \qquad (6)$$

where $p_k(j)$ is the probability that the S will read j when pointer position k was presented. The Response Equivocation is much larger with unlimited responses, but is also more constant from series to series.

With the method of computation used here, the Information Transmitted is found by subtracting the Response Equivocation from the Response Information. These amounts are given in lines four and eight show that the amount of information transmitted to the *S*s is about the same for both conditions of response limitation. It is important to see, however, that the conditions under which the information is transmitted under the two instructions are quite different. With Limited Responses, the transfer remains as efficient as it does because the equivocation is minimized. With Unlimited Responses the transfer of information occurs in spite of considerable equivocation because of the use by the *S*s of a larger number of response categories.

In the introduction we stated that the amount of information transmitted should be invariant provided that the number of pointer positions used is equal to or larger than some minimum number. This statement is supported by the data of Table II, which show that the amount of information transmitted under both response limitations was essentially constant when ten or more pointer positions were used. When only five positions were permitted the amount transmitted was depressed, of course, because of the limitation imposed by the smallness of I_s.

The maximum information which can be transmitted under the conditions of this experiment is a little more than three bits for every appearance of the pointer. This means that no advantage in information transmission is gained by the use of more than about ten different positions. This effect, of course, can be complicated somewhat by the number habits of the *S*s.

Discussion

The error analysis showed that the error involved in the transmission of information about a continuous variable is not increased when this variable is represented by discrete pointer positions which, in effect, "round-off" the original values, provided that ten or more positions are used and that the *S*s are required to read only values which correspond exactly to the presented positions. If the *S*s are unaware of the pointer restrictions, then the error is increased. The information analysis, on the other hand, showed that the amount of information transmitted is approximately constant whether or not the *S*s were aware of the pointer restrictions.

It is apparent that the two analyses do not reveal the same facts about the errors of transmission. This can be seen most clearly in the case of the Unlimited Response condition. The *S*s under this condition responded to each pointer position with a wide variety of different values, and this greater variety of response resulted in a large Observer Error which inflated the Total Error. The greater variety of response did not decrease the amount of Information Transmitted, however, because the distribution of responses to a given position did not seriously overlap the distribution of any other position. Little or no overlap among these response distributions means that, given a particular response, we can make a statement with considerable assurance about which pointer position has occurred.

The information analysis responds only to the coding possibilities in an information transmission system. The error analysis does not have this advantage since a greater variety of response shows up only as increased error whether or not the overlap among the response distributions exists. On the other hand, the information analysis has the disadvantage that it is not sensitive to magnitude of errors. Thus, if two different situations show that the same amount of information has been transmitted, we cannot conclude that the two situations will give the same magnitude of errors.

In order to demonstrate this point, we have determined the frequencies with which errors of various magnitudes were made in the situations of this experiment. Figure 1 shows the proportions of times that errors of various magnitudes were not exceeded when Limited Responses were used for the four different numbers of permissible pointer positions. These curves are constructed with the following rationale: Suppose that only five pointer positions were used, and that position "20" was presented. If *S* responded with "20," his maximum error would be ten units, since the position "20" represented values between 10 and 30. If *S* read that position as "40," then his maximum error would be 30 units. A similar argument can be used for other numbers of pointer positions. The plotted points are connected by straight lines on the assumption that all values of X occur equally often.

When only five pointer positions

were used, errors of larger magnitude were made quite frequently. The Ss rarely made an error in reporting the presented position, but the coding was so coarse that many errors with respect to the original variable occurred. For the other three numbers of pointer positions, errors of different magnitudes were made about equally often, although the 10-position case gave fewer errors of intermediate magnitude.

Thus, when responses were limited to pointer position values, the uniformity in amount of information transmitted indicated that the same number of errors of different magnitudes occurred when the three larger numbers of pointer positions were used. When Unlimited Responses were used, however, the magnitude of errors was increased considerably. Figure 2 shows the proportion of times errors of various magnitudes were made when just ten pointer positions were used with the two conditions of response limitation. It is clear that the use of Unlimited Responses greatly increased the frequency with which errors of any given magnitude were exceeded. This increased error results from the following factors: If only ten pointer positions were used, and position "50" were presented, the maximum error possible when the reading is "50" is five units. When Unlimited Responses are used, however, *S* may read it as "54," in which case the maximum error with respect to the original *X* variable is nine units. Such errors can occur, even though there is no overlap of the distributions of readings used for the various presented positions.

If we rounded off the readings made by *S* with Unlimited Responses into

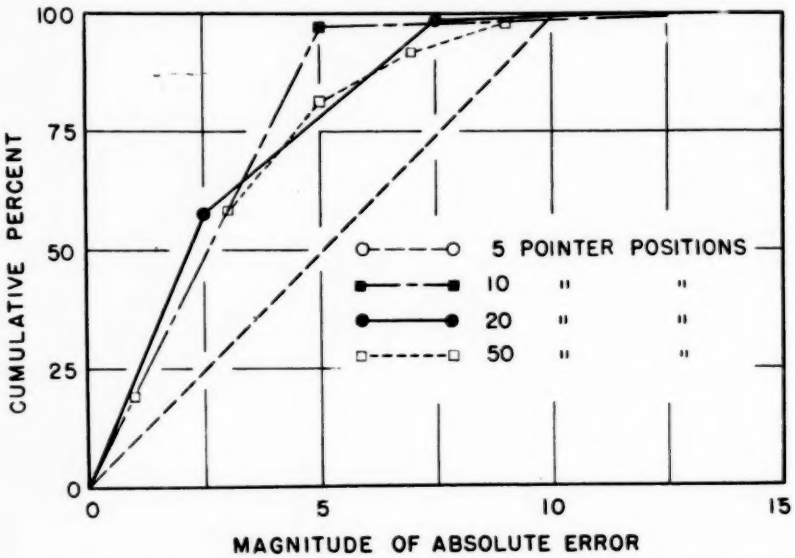

FIG. 1. The per cent of errors which are as small as the magnitude indicated on the abscissa. Each curve is for a different number of possible pointer positions, and all curves were obtained with Limited Responses.

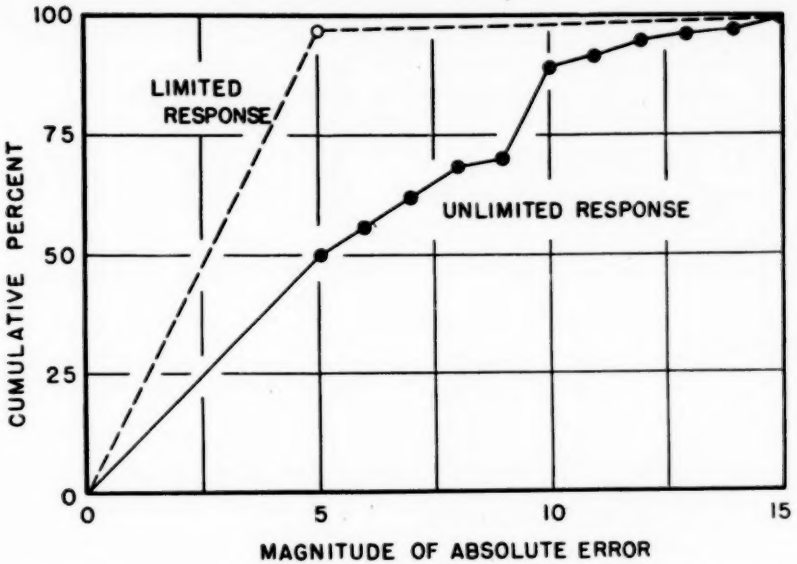

FIG. 2. The per cent of errors which are as small as the magnitude indicated on the abscissa. Each curve is for a different condition of response limitation, and both curves were obtained with ten possible pointer positions.

groups whose mid-points are the same as the presented values, a reading of "54" to a presented position of "50" would be considered a correct reading. Yet, because he was allowed to use that number, his actual error is larger. Thus, the number of correct readings (as determined by our correct prediction of which discrete pointer position occurred, knowing the S's reading of an unknown pointer position) might be the same for both cases, but the magnitudes of errors larger. The amount of information in bits indicates something only about the number of readings which are correct in this sense, not about the magnitude of an error. It does indicate, however, the minimum number of discrete pointer positions from which the maximum precision can be obtained.

SUMMARY

This experiment studied the accuracy with which *Ss* can determine the value of a continuous variable when only discrete values representing ranges of the continuous variable are presented. The *Ss* made readings on an instrument dial when 5, 10, 20, or 50 different positions of a pointer were possible within a single interpolation interval, under instruction to restrict their responses to values represented by the discrete positions and under instructions to use many different possible readings. The results and conclusions are:

1. When *Ss* were limited in their responses to values corresponding to the presented values, the total mean square error (including the rounding-off, or instrument, error)

was essentially constant for 10, 20, or 50 pointer positions. When *Ss* were allowed Unlimited Responses, the total mean square error increased with fewer numbers of different discrete pointer positions.

2. An analysis of the amount of information transmitted by the discrete pointer positions showed that the same amount of information was transmitted with either Limited or Unlimited Responses, and for all numbers of pointer positions greater than ten.

3. The amount of information measures the extent to which responses to the various discrete stimuli do not overlap, but is not a measure of the magnitude of errors. Thus, the fact that the same amount of information was transmitted with either Limited or Unlimited Responses does not indicate that the magnitude of errors is the same for both cases. The measure of information does indicate, however, the minimum number of discrete pointer positions from which the maximum precision can be obtained.

(Manuscript received for priority publication September 17, 1951)

References

1. Chapanis, A. Theory and methods for analyzing errors in man-machine systems. *Ann. N. Y. Acad. Sc.*, 1951, **51**, 1179–1203.
2. Chapanis, A., Garner, W. R., and Morgan, C. T. *Applied experimental psychology.* New York: Wiley, 1949.
3. Garner, W. R., and Hake, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, **58**, 446–459.
4. Kappauf, W. E. The design and use of instruments. In *Human factors in undersea warfare.* Washington, D. C.: Committee on Undersea Warfare, National Research Council, 1949.

Classics Among

PSYCHOLOGICAL MONOGRAPHS

Thorndike, E. L. The Mental Life of the Monkey. 1899, #15. *$.50*

Carr, Harvey. Visual Illusion of Movement During Eye Closure. 1905, #31. *$1.25*

Watson, John B. Kinaesthetic and Organic Sensations: Their Role in the Reactions of the White Rat to the Maze. 1907, #33, *$1.00*

Shepherd, William T. Some Mental Processes of the Rhesus Monkey. 1910, #52. *$.75*

Franz, Shepard Ivory and Lafora, Gonzalo R. On the Functions of the Cerebrum: The Occipital Lobes. 1911, #56. *$1.25*

Fernberger, Samuel W. On the Relation of the Methods of Just Perceptible Differences and Constant Stimuli. 1912, #61. *$1.00*

Boring, Edwin G. Learning in Dementia Praecox. 1913, #63. *$1.00*

Langfeld, Herbert S. On the Psychophysiology of a Prolonged Fast. 1914, #71. *$.75*

Franz, Shepard Ivory. I. Symptomological Differences Associated with Similar Cerebral Lesions in the Insane. II. Variations in Distribution of the Motor Centers. 1915, #81. *$1.50*

Peckstein, Louis Augustus. Whole versus Part Methods in Motor Learning. 1917, #99. *$.75*

Kjerstad, Conrad L. The Form of the Learning Curves for Memory. 1919, #116. *$1.25*

Tolman, Edward C. Retroactive Inhibition as Affected by Conditions of Learning. 1918, #107. *$.75*

MANY OF THE EARLY MONOGRAPHS ARE OUT-OF-PRINT.
ONLY A LIMITED QUANTITY OF THE ABOVE
NUMBERS ARE AVAILABLE.

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Ave. N.W., Washington 5, D. C.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The issues listed below are available for sale

Year	Volume	Available Numbers						Price per Number	Price per Volume
1916	1	1	2	3	4	5	6	$1.25	$7.00
1917	2	1	2	–	4	5	–	$1.25	$5.00
1920	3	–	2	3	4	5	6	$1.25	$6.25
1921	4	1	2	3	4	5	6	$1.25	$7.00
1922	5	–	–	3	4	5	6	$1.25	$5.00
1923	6	–	–	–	–	–	–	——	——
1924	7	1	2	–	–	5	–	$1.25	$3.75
1925	8	1	2	3	4	–	6	$1.25	$6.25
1926	9	–	2	3	–	5	6	$1.25	$5.00
1927	10	–	–	3	4	5	6	$1.25	$5.00
1928	11	1	2	3	4	5	6	$1.25	$7.00
1929	12	1	2	3	4	5	6	$1.25	$7.00
1930	13	1	2	3	4	5	6	$1.25	$7.00
1931	14	1	2	3	4	5	6	$1.25	$7.00
1932	15	1	2	3	4	5	6	$1.25	$7.00
1933	16	1	2	3	4	5	6	$1.25	$7.00
1934	17	1	2	3	4	5	6	$1.25	$7.00
1935	18	1	2	3	4	5	6	$1.25	$7.00
1936	19	1	2	3	4	5	6	$1.25	$7.00
1937	20	1	2	3	4	5	6	$1.25	$7.00
	21	1	2	3	4	5	6	$1.25	$7.00
1938	22	1	2	3	4	5	6	$1.25	$7.00
	23	1	2	3	4	5	6	$1.25	$7.00
1939	24	1	2	3	4	5	6	$1.25	$7.00
	25	1	2	3	4	5	6	$1.25	$7.00
1940	26	1	2	3	4	5	6	$1.25	$7.00
	27	1	2	3	4	5	6	$1.25	$7.00
1941	28	1	2	3	4	5	6	$1.25	$7.00
	29	1	2	3	4	5	6	$1.25	$7.00
1942	30	1	2	3	4	5	6	$1.25	$7.00
	31	1	2	3	4	5	6	$1.25	$7.00
1943	32	1	2	3	4	5	6	$1.25	$7.00
	33	1	2	3	4	5	6	$1.25	$7.00
1944	34	1	2	3	4	5	6	$1.25	$7.00
1945	35	1	2	3	4	5	6	$1.25	$7.00
1946	36	1	2	3	4	5	6	$1.25	$7.00
1947	37	1	2	3	4	5	6	$1.25	$7.00
1948	38	1	2	3	4	5	6	$1.25	$7.00
1949	39	1	2	3	4	5	6	$1.25	$7.00
1950	40	1	2	3	4	5	6	$1.25	$7.00
1951	41	By subscription, $14.00 (2 vols.)						$1.25	——
	42								
		List price, Volumes 1 through 40							$260.25
		30% discount							78.08
		Net price, Volumes 1 through 40							$182.17

Information about the Journal of Experimental Psychology: as noted in the table, the journal was not published in 1918 and 1919; from 1937 to 1943 it was published at the rate of two volumes per year; from 1944 to 1950, in one volume per year. Beginning in 1951 there are two volumes a year.

Information about prices: when all numbers are available, the price per volume throughout the series is $7.00. For incomplete volumes, the price is $1.25 for each available number. For foreign postage, $.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of $ 50.00 and over
20% on orders of $100.00 and over
30% on orders of $150.00 and over

Address orders and subscriptions to

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Avenue N.W. Washington 5, D. C.